

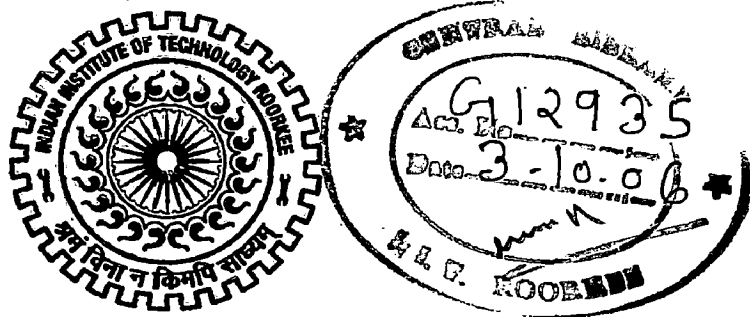
# **HYDROLOGIC RESPONSE OF KANKAIMAI WATERSHED IN EASTERN NEPAL**

**A DISSERTATION**

*Submitted in partial fulfillment of the  
requirements for the award of the degree  
of*  
**MASTER OF TECHNOLOGY**  
*in*  
**WATER RESOURCES DEVELOPMENT**

**By**

**MEKH NATH SHARMA**



**DEPARTMENT OF WATER RESOURCES DEVELOPMENT AND MANAGEMENT  
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE  
ROORKEE -247 667 (INDIA)  
JUNE, 2006**

## CANDIDATE'S DECLARATION

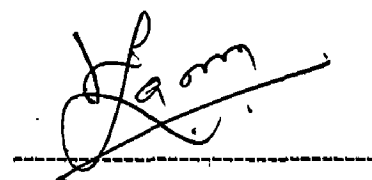
---

I do hereby declare that the dissertation entitled, "**HYDROLOGIC RESPONSE OF KANKAIMAI WATERSHED IN EASTERN NEPAL**" is being submitted by me in partial fulfillment of requirement for the award of degree of Master of Technology in Water Resources Development and submitted at the Department of Water Resource Development and Management, Indian Institute of Technology, Roorkee is an authentic record of my own work carried out during the period from July 2005 to June 2006 under the guidance of Dr. Nayan Sharma, Professor, Department of Water Resource Development and Management and Dr. Zulfequar Ahmad, Assistant Professor, Department of Civil Engineering, Indian Institute of Technology Roorkee.

The matter embodied in the dissertation has not been submitted by me for the award of any other degree.

Dated: 28<sup>th</sup> June, 2006

Place: IIT Roorkee



(Mekh Nath Sharma)

WRD&M, IIT Roorkee

## CERTIFICATE

---

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.



Dr. Zulfequar Ahmad

Assistant Professor

CED, IIT Roorkee



Dr. Nayan Sharma

Professor

WRD&M, IIT Roorkee

## ACKNOWLEDGEMENT

---

I feel privileged to express my deep sense of gratitude and sincere regards to Dr. Nayan Sharma, Professor, Department of Water Resource Development and Management and Dr. Zulfequar Ahmad, Assistant Professor, Department of Civil Engineering, Indian Institute of Technology Roorkee for their keen interest, invaluable guidance and constant encouragement throughout the course of present study.

I am pleased to extend my sincere gratitude to Prof. S.K. Tripathi, Prof. G. Chauhan, Prof. G.C. Mishra, Prof. U.C. Chaube, Prof. B.N. Asthana, Prof. R.P. Singh, Prof.D.D.Das, Dr.Deepak Khare, Dr.M.L. Kansal and Dr.S.K. Mishra of WRD&M, IIT Roorkee for their valuable suggestions and encouragement.

I am also thankful to all the staffs of WRD&M for their kind cooperation and support during my stay here.

On this occasion, I would like to share my happiness with all my fellow trainee officers, and extend my deep gratitude for creating warm and homely atmosphere throughout the period of course.

I would also like to express my thanks to Mrs Kanthi Chandralata for her support during the course of this study.

I also like to acknowledge the government of India for the scholarship granted to me to pursue M-Tech. Course.

My wife Archana and daughters Niharika and Prasiddika deserve very special acknowledgement for their continuous support, passion and encouragement throughout the course.

## ABSTRACT

---

Soil erosion is a serious environmental problem in Nepal where more than 80 percent of the land area is mountainous and tectonically active. Anthropogenic causes such as deforestation, overgrazing and intensive farming have accelerated the erosion problem. Unscientific cultivation, haphazard construction and intensive monsoon have further aggravated the situation. Thus, it is very important to understand the erosion process and assess the magnitude of the problem so that effective counter measures and appropriate sediment management method can be implemented.

The rapid growth in population, urbanization and industrialization, and economic and social changes during last few years has resulted in an increased and diversified demand of water. At the same time, quantity of available water from surface and ground water sources has remained constant. So water has progressively emerged to become one of the most pressing national issues of present time on the development and environment.

Thus there is need to use recent advancement in technology and modeling tool to develop optimal water resources system and manage sediment problems. Rainfall-runoff and sediment yield models are the key components of the overall modeling framework for sustainable watershed development and management. The process of runoff and sediment erosion is influenced by the spatial controls exerted by the land surface such as elevation, slope, drainage network and vegetation cover. So the knowledge of geomorphology and vegetation of watershed can enhance one's understanding and capability to model the processes.

The present study has been taken up in the above background to develop rainfall-runoff and sediment yield models to predict runoff and sediment yield from Kankaimai watershed in eastern Nepal. Geomorphologic and vegetational analysis was carried out using remote sensing and geographic information system. The study has revealed that the Kankaimai watershed is fairly good with moderately high peak flow of shorter duration giving quick response of sediment yield and runoff.

With this knowledge of watershed characteristic and statistical analysis, eight different combinations of daily rainfall, runoff and temperature for runoff models and four combinations of daily runoff and sediment yield for sediment yield prediction models were selected as input data. Artificial Neural Network (ANN) models were developed using selected combinations of data as input to a three layered back-propagation feed-forward neural network. The regression models were also developed using all above combinations of input data and compared with the results obtained from ANN models. Nash coefficient ( $R^2$ ), coefficient of correlation (CC), root mean square error (RMSE) and recovery ratio ( $R_R$ ) were estimated to ascertain the model performance. ANN model validation statistics resulted in  $R^2$  values ranging from 0.35 to 0.82, CC values from 0.60 to 0.91, RMSE values from 192.29 m<sup>3</sup>/s to 103.67 m<sup>3</sup>/s and  $R_R$  from 0.59 to 0.96 for runoff prediction and  $R^2$  values from 0.80 to 0.93, CC values from 0.90 to 0.97, RMSE values from 19058 t/day to 34345 t/day and  $R_R$  values from 0.93 to 1.18 for sediment yield prediction. Using regression models,  $R^2$  values for the same data set varied from 0.10 to 0.60 ( $0.81 > CC > 0.384$ ,  $226.33 > RMSE > 157.45$ ,  $0.758 > R_R > 0.614$ ) for runoff prediction and  $-0.85$  to  $-0.16$  ( $CC=0.97$ ,  $106607 > RMSE > 83770$ ,  $1.65 > R_R > 1.46$ ) for sediment yield prediction. Performance evaluation of different models suggests that antecedent runoffs of time step 't-1' and 't-2' as an additional inputs variable alongwith concurrent rainfall improve the performance of runoff model for this watershed. Similarly the concurrent runoff has very high correlation with sediment yield. Use of antecedent runoff and sediment yield as additional input variables to concurrent runoff does not improve the model performance. Furthermore ANN model performed better than the regression equations.

# CONTENTS

---

	DESCRIPTION	PAGE NO.
	Candidate's Declaration	i
	Acknowledgement	ii
	Abstract	iii
	Contents	v
	List of Tables	viii
	List of Figures	ix
	List of Symbols and Abbreviations	xi
<b>CHAPTER - 1</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	Background	1
1.2	Brief Review	1
1.3	Sediment Yield	3
1.4	Objectives	4
1.5	Scope of works	5
<b>CHAPTER- 2</b>	<b>LITERATURE REVIEW</b>	<b>6</b>
2.1	Watersheds Geomorphology	6
2.2	Normalized Difference Vegetation index	7
2.3	Hydrologic Modeling	8
2.3.1	Empirical Models	9
2.3.2	Physically Based Models	12
2.3.3	Neuromorphic Models	14

<b>CHAPTER – 3</b>	<b>DESCRIPTION OF THE STUDY AREA</b>	<b>18</b>
3.1	General	18
3.2	River Characteristics	21
3.3	Hydrological and Meteorological Data	21
	(a) Climate	21
	(b) Rainfall	22
	(c) Runoff	23
	(d) Sediment Yield Data	24
3.4	Land Resources	24
3.5	Existing Water Resources Utilization	25
<b>CHAPTER- 4</b>	<b>GEOMORPHOLOGIC &amp; VEGETATIONAL ANALYSES</b>	<b>27</b>
4.1	Background	27
4.2	Geomorphologic Analysis	27
4.3	Vegetational Analysis	35
<b>CHAPTER – 5</b>	<b>ANN MODEL DEVELOPMENT</b>	<b>38</b>
5.1	Introduction of Artificial Neural Network	38
	5.1.1 General	38
	5.1.2 Mathematical Aspects	40
	5.1.3 Network Training	41
	5.1.4 Back-propagation	42
	5.1.5 Advantage of ANN	43
	5.1.6 Limitation of ANN	43
5.2	Development of ANN Models	44
	5.2.1 Data Partition	44

5.2.2	Selection of Input and Out Variables and their Normalization	45
5.2.3	Designing an ANN	48
5.2.4	Performance Evaluation Criteria of Models	49
5.2.5	Rainfall-Runoff Models	50
5.2.6	Sediment Yield Prediction Models	56
<b>CHAPTER- 6</b>	<b>DEVELOPMENT OF REGRESSION MODEL</b>	<b>59</b>
6.1	Introduction	59
6.2	Regression Equation	59
6.3	Rainfall-Runoff Models	60
6.4	Sediment Yield Models	63
<b>CHAPTER- 7</b>	<b>RESULTS AND DISCUSSIONS</b>	<b>65</b>
7.1	Geomorphologic and Vegetation Analysis	65
7.1.1	Geomorphologic Analysis	65
7.1.2	Vegetational Analysis	66
7.2	ANN Models	67
7.2.1	Rainfall-Runoff Models	67
7.2.2	Sediment yield prediction Models	69
7.3	Regression Models	71
7.4	Comparison of ANN and Regression models	71
<b>CHAPTER - 8</b>	<b>CONCLUSIONS</b>	<b>75</b>
	<b>REFERENCES</b>	<b>77</b>
	<b>APPENDIX – A DAILY TEMPERATURE DATA</b>	<b>82</b>
	<b>APPENDIX - B DAILY RAINFALL DATA</b>	<b>90</b>
	<b>APPENDIX - C DAILY RUNOFF DATA</b>	<b>98</b>
	<b>APPENDIX - D DAILY SEDIMENT YIELD DATA</b>	<b>101</b>



## **LIST OF TABLES**

---

Table 3.1	Mean Monthly Temperature, Relative Humidity and Sunshine Hours of Kankaimai Basin .....	21
Table 3.2	Mean Monthly Rainfall of Stations in and around Kankaimai Basin.....	22
Table 3.3	Stream Gauge Stations at and around Kankaimai Basin.....	23
Table 3.4	Mean Monthly Runoff of Kankaimai Basin and Bub-basin.....	23
Table 3.5	Monthly Suspended Sediment Yield .....	24
Table 4.1	Details of Streams of Different Orders and Their Length.....	30
Table 4.2	Land Coverage and NDVI values.....	37
Table 4.3	Land Coverage in Percentage .....	37
Table 5.1	Selected Input and Output Variables.....	45
Table 5.2	Runoff Prediction Models .....	54
Table 5.3	Sediment Yield Prediction Models .....	54
Table 5.4	Details of Weighting Factors and Bias of Runoff Prediction Model .....	55
Table 5.5	Details of Weighting Factors and Bias of Sediment Yield Prediction Model.	58
Table 6.1	Performance of Various Regression Models for Runoff Prediction.....	61
Table 6.2	Performance of Various Regression Models for Sediment Yield Prediction..	61
Table 7.1	Summary of Geomorphologic Parameters.....	66

## LIST OF FIGURES

---

Fig. 2.1	Reflectance Characteristics of Various Objects with Wavelength.....	7
Fig. 3.1	Location Map of Study Area .....	19
Fig. 3.2	Study Area Map.....	20
Fig. 4.1	Strahler River Ordering System.....	28
Fig. 4.2	Relationship between Stream Order and Stream Number .....	30
Fig. 4.3	Relationship between Stream Order and Stream Length.....	31
Fig. 4.4	Land Cover Map .....	36
Fig. 5.1	Configuration of Feed Forward Three Layers ANN .....	39
Fig. 5.2	Schematic Diagram of Typical $j^{\text{th}}$ Node.....	40
Fig. 5.3	Autocorrelation of Runoff and Cross Correlation between Runoff and Rainfall.....	47
Fig. 5.4	Cross Correlation between Runoff and Temperature .....	47
Fig. 5.5	Autocorrelation of Sediment Yield and Cross Correlation between Sediment Yield and Runoff .....	48
Fig. 5.6	ANN Architecture of Runoff Prediction Model 4 .....	51
Fig. 5.7	Validation of Computed Runoff with Observed one .....	52
Fig. 5.8	Mass Curve of Computed Runoff and Observed Runoff for Validation Data.....	52
Fig. 5.9	Scatter Plot of Computed Runoff against Observed Runoff for Validation Data.....	53
Fig. 5.10	ANN Architecture of Sediment Yield Prediction Model 1.....	56
Fig 5.11	Validation of Computed Sediment Yield with Observed one.....	57

Fig. 5.12	Mass Curve Plot of Computed Sediment Yield and Observed Sediment Yield for Validation data.....	57
Fig. 5.13	Scatter Plot of Computed Sediment Yield against Observed Sediment Yield.....	58
Fig. 6.1	Validation of Computed Runoff with Observed one.....	62
Fig. 6.2	Mass curve of computed runoff and observed runoff for validation data .....	62
Fig. 6.3	Scatter Plot of Computed Runoff against Observed Runoff for Validation Data .....	63
Fig. 6.4	Validation of Computed Sediment Yield with Observed one.....	63
Fig. 6.5	Mass Curve of Computed Sediment Yield and Observed Sediment Yield for Validation Data .....	64
Fig. 6.6	Scatter Plot of Computed Sediment Yield against Observed Sediment Yield for Validation Data.....	64
Fig. 7.1	Comparison of ANN and Regression Runoff Prediction Models.....	73
Fig. 7.2	Comparison of ANN and Regression Sediment Yield Prediction Models ....	73
Fig. 7.3	Comparison of ANN Regression Models for Runoff Prediction (Mass Curve).....	74
Fig. 7.4	Comparison of ANN Regression Models for Sediment Yield Prediction (mass curve).....	74

## LIST OF SYMBOLS AND ABBREVIATIONS

---

### Symbols:

$A_w$	Total watershed area
$A_c$	Area of Circle
$b$	Bias
$B$	Bias vector
$C$	Crop management factor
$C_c$	Constant of channel maintenance
$CC$	Coefficient of correlation
$C_f$	Channel segment frequency
$D$	Diameter of circle
$D_d$	Drainage density
$D_F$	Drainage factor
$E$	Kinetic energy of rainfall
$E_f$	Error function
$F$	Estimated rate of soil erosion
$G$	Transport capacity of overland Flow
$H_{max}$	Maximum elevation of watershed
$H_{min}$	Minimum elevation of watershed
$K$	Soil detachability index
$L$	Length of streams
$L_b$	Longest length of watershed
$L_c$	Length of drainage line
$L_p$	Perimeter of basin
$n$	Number of pass or epoch
$N$	Number of streams
$p$	Number of output nodes
$P_t$	Number of training pattern
$P$	Support practice
$R^2$	Nash Coefficient
$Q$	Runoff
$Q_s$	Sediment yield

R	Rainfall
$R_A$	Area ratio
$R_B$	Bifurcation ratio
$R_c$	Circularity ratio
$R_e$	Elongation ratio
$R_h$	Relief ratio
$R_L$	Channel length ratio
$R_p$	Relative relief
$R_R$	Recovery ratio
S	Sine of slope gradient
$S_w$	Slope of watershed
t	Time
T	Temperature
$T_c$	Time of concentration
$w_s$	Order of streams
w	Connection weight
W	Weight vector
x	Input variable
X	Input vector
$\hat{y}$	Target variables
y	Output variable
Y	Output vector
$\hat{Y}$	Target vector
Z	Percentage of validation data falling within 50% error line
$\beta$	Slope parameter
$\Delta w$	Weight increments
$\epsilon$	Learning rate
$\alpha$	Momentum factor

#### Abbreviations:

AI	Artificial Intelligence
ANGPS	Agriculture No-Point Source Model

ANN	Artificial Neural Network
ANSWERS	Arial Non-point Source Watershed Environment Response Simulation
CREAMS	Chemical Runoff and Erosion from Agricultural Management Systems
DEM	Digital Elevation model
DHM	Department of Hydrology and Meteorology
DOI	Department of Irrigation
FL	Fuzzy Logic
FMIS	Farmers Managed Irrigation System
GA	Genetic Algorithm
GCP	Ground Control Point
GIS	Geographic Information System
HSPF	Hydrologic Simulation Package-Fortran
IHDM	Institute of Hydrology Distributed Model
ISP	Irrigation Sector Project
LS	Slope Length
MHP	Mechi Hill Project
MLP	Multi Layer Perceptron
MMF	Morgan Morgan and Finney
MUSLE	Modified Universal Soil Loss Equation
NCDC	Namsaling Community Development Centre
NDVI	Normalized Difference Vegetation Index
NEA	Nepal Electricity Authority
NIR	Near Infra Red
NWS	National Weather Service
RMSE	Root Mean Square Error
RUSLE	Revised Universal Soil Loss Equation
SCS	Soil Conservation Service
SHE	System Hydrologique European
SISP	Second Irrigation Sector Project
SSARR	Streamflow Synthesis and Reservoir Regulation
SWM	Stanford Watershed Model
SWMM	Storm Water Management Model
TM	Thematic Mapper
USDA	U.S. Department of Agriculture
USLE	Universal Soil Loss Equation
UTM	Universal Transverse Mercator
VDC	Village Development Committee

VIS

Visible

WECS

Water Energy Commission Secretariat

WEPP

Water Erosion Prediction Project

# CHAPTER- 1

## INTRODUCTION

---

### 1.1 Background

A watershed is a land area which drains into a stream system, upstream from its mouth or other designated point of interest. Surface characteristic, soil depth, geological structures, topography and climate of the watershed play an interrelated role in the behavior of water, which flows over or through it. Watersheds are subjected to many types of modifications by human and natural activities. Such changes can be distinguished as point changes and non-point changes and affect virtually all elements of hydrologic cycle. Structural changes such as dam construction, channel improvement, detention storage etc. are examples of point changes. Forestry, agriculture, mining, urbanization etc. are non-point land use changes. There has been a growing need to study, understand and quantify the impact of major land use changes on hydrologic regime, both water quantity and quality.

Hydrological modelling is a powerful technique for investigation of hydrologic system for both hydrologists and the practising water resources engineers involved in the planning, development and management of water resources. These models are used for, varied purposes such as planning and designing soil conservation practices, irrigation water management, wetland restoration and water table management .On a large scale, models are used for flood protection projects, rehabilitation of aging dams, flood plain management, water quality evaluation, and water supply forecasting.

### 1.2 Brief Review

The early history of hydrology and hydraulic structures begins thousands of years ago when archeological evidence suggests that a dam was built across the Nile as early as 4000 B.C. Later a canal to transport fresh water was constructed between Cairo and Suez (Bedient and Huber, 1992). In the seventeenth century, Perrault made the first recorded measurements of rainfall and surface flow (Ryan, 1999).



It was not until relatively recently that attempts were made to model rainfall-runoff thus predicting runoff hydrographs, peak flow rates, and times to peak. Early models based on empirical equations predicted peak discharge and time to peak. In 1932, Sherman proposed the “unitgraph” or unit hydrograph technique. It was one of the first attempts to predict an entire hydrograph instead of just the peak flow and time to peak (Kilgore, 1997). Many researchers followed with increasingly complex models to improve the unit hydrograph shape. Although these techniques produced mathematically correct hydrographs, Todini (1988) states that their connections with the “real world” were lost. During the 1960’s and 1970’s, researchers focused their efforts on developing models with parameters having a physical interpretation. Due to limitations in the amount of available data and computing power, these physically based parameters were aggregated or lumped together, thus greatly decreasing the amount of data to be processed. These models with aggregated parameters are termed conceptual lumped parameter models. The rapid increase in computing power of the 1980’s and 1990’s has brought more complex physically based distributed models. Parameters no longer need to be lumped together because of computing limitations.

Physically based Distributed parameter models are capable of incorporating information about the spatial variability of soils, land use, topography, or any other parameter in the modeling scenario. The improved availability of geographic information systems (GISs) aids in managing the large amounts of data required for distributed parameter models. GIS software can be combined with digital data such as soil type, vegetative cover, land use, and digital elevation models (DEMs) to create hydrologic models or input to hydrologic models.

Physically based distributed parameter models / conceptual lumped parameter models attempts to capture the characteristic of underlying physical processes through the use of equations of mass, momentum, and energy in the case of deterministic models and their simplified forms in the case of conceptual models. Though physically based deterministic hydrological models have proved to be very useful for simulation of various processes related to the management of water such as hydrodynamic, morphological, ecological, water quality, sediment yield, groundwater flow etc., implementation and calibration of such a models can typically present various difficulties, requiring sophisticated mathematical tools (Sorooshian et al, 1993), significant amounts of calibration data (Yapo et al, 1996), some degree of expertise and experience with models (Kisi, 2005).

Thus there is a need to investigate an alternative approaches to model the hydrologic responses of a watershed. Soft computing technique is one the alternative approaches to deal with such problems. Artificial neural network (ANN) is one of the soft computing techniques which is composed of densely interconnected processing nodes and has the ability to extract and store the information from the few patterns (data) in training through learning. The model is easy to develop; yields satisfactory results when applied to complex systems. Hydrologic applications of ANN include the modeling of rainfall-runoff forecasting, sediment yield process, snow-rainfall process, assessment of stream's ecological and hydrological responses to climate change, and ground water quality prediction and ground water remediation.

In the hydrological forecasting context, recent experiment has reported that ANNs may offer a promising alternative for Rainfall- Runoff modeling (Tokar and Johnson, 1999). Similarly neural network (NN) approaches to model the streamflow-suspended sediment relationship were investigated by Kisi (2005) and Agrawal ( 2005). In many or most occasions it was shown that the neural networks tend to give better result than the deterministic models, provided that the process under consideration is not changed in time.

### **1.3 Sediment yield**

Soil degradation by accelerated water and wind-induced erosion is a serious problem and will remain so during the 21st century, especially in developing countries of tropics and subtropics. Erosion is a natural geomorphic process occurring continually over the earth's surface. However, the acceleration of this process through anthropogenic perturbations can have severe impacts on soil and environmental quality.

Soil erosion and it off-site downstream damages are major concern around the world causing loses in soil productivity and degradation of landscape. It has been estimated that human induced soil degradation affected 15% of world arable land surface. Estimate of global soil erosion rates ranges from .88 mm/yr to .30mm /year from land surface that is carried downstream to lakes, reservoirs and estuaries leading to reduction in their storage capacities and affecting water quality, navigation, and biological productivity.

The total land area subjected to human-induced soil degradation is estimated at about 2 billion ha (Lal, 2001). Of this, the land area affected by soil degradation due to erosion is estimated at 1100 Mha by water erosion and 550 Mha by wind erosion. South Asia is one of the regions in the world where soil erosion by water and wind is a severe problem.

In Nepal, over the last century, the proportion of the erosion induced by growing population pressures on a limited land base has increased considerably. Human factors such as unscientific cultivation, destruction of natural vegetation, improper land use and haphazard construction, together with natural factors such as fragile geology, steep slopes, and intensive monsoon rains has exacerbated erosion and resulted in the environmental degradation of the country. It has been estimated that as much as 1.63 mm of topsoil are displaced from the total land surface of Nepal every year. Hence there is a need to understand the soil erosion process and quantify its magnitude so that soil erosion management plan in the watershed and sediment control measures in the downstream reach could be developed.

In such a situation it becomes necessary to understand the behaviour of watershed in terms of its runoff and sediment yield process and develop suitable models which can yield sufficiently accurate results with shorter length of data . In the present study also, attempt has been made to simulate the runoff and sediment yield process of Kankaimai watershed in the eastern Nepal by using Artificial Neural Networks.

#### **1.4 Objectives**

In the light of above review related to rainfall-runoff and sediment yield models, the following objectives were considered in the present study .

- a. Formulation of rainfall-runoff models for Kankaimai watershed in eastern Nepal using ANN.
- b. Establishment of correlation between river discharge and sediment yield at a Mainachuli station of Kankaimai river in eastern Nepal using ANN.

## **1.5 Scope of works**

The scopes of proposed research are as follows:

- Review of past studies and literature relating to the proposed research .
- Geomorphologic studies of the watershed
- Vegetational analysis of the watershed
- Selection of different input data through statistical procedure
- Investifation of different ANN models for runoff prediction .
- Investigation of different ANN models for sediment yield prediction .
- Evaluation of results and selection of appropriate models.
- Comparison of ANN models with linear regression models.

## CHAPTER-2

### LITERATURE REVIEW

---

#### 2.1 Watershed geomorphology

The importance of geomorphology for hydrological purposes has been increasingly appreciated among hydrologists in the last few decades when they became engaged in the development of water resources in developing and other countries where the available hydrological data were found to be inadequate for their aims. Observation on discharge, sediment load, etc, over a sufficient period of time, are often few in number or totally absent. Thus hydrologist was obliged to use different unconventional approaches to evaluate the characteristics of rivers and drainage basins and to get a proper idea of the order of magnitude of various aspects of the water problems (Kuiper).

The study of fluvial geomorphology and morphometrical and environmental geomorphological analysis of drainage basin have become increasingly important. In fact geomorphology was found to have very close links with both surface and sub surface water condition. Schumm(1964), emphasized the role of geomorphology and mentioned that a general relationship exists between hydrological geomorphological variables. Once these relation are established, the probable hydrological characteristics of other areas, which are geomorphologically similar, can be estimated (Verstappen,1983).

The drainage basin, or watershed, is the fundamental unit in geomorphology within which may be studied the relations between landforms and the processes that modify them. The study of basin morphometry attempts to relate basin and stream network geometries to the transmission of water and sediment through the basin. The size of a drainage basin influences the amount of water yield; the length, shape, and relief affect the rate at which water is discharged from the basin and the total yield of sediment; the length and character of the streams channels affect the availability of sediment for stream transport and the rate at which water and sediment are discharged. Drainage density, constant of channel Maintenance, stream order, elongation ratio, circulatory ratio, form factor, bifurcation ratio, relief ratio and relative ratio are some of the important parameters that need to be computed so that

characteristics of rivers and drainage basin could be evaluated . Details of these parameters are discussed in chapter 4 of section 4.2.

## 2.2 Normalized Difference Vegetation Index (NDVI)

The normalized difference of the vegetation index [NDVI] is a non-linear transformation of the visible (red) and near-infrared bands of satellite information. NDVI is defined as the difference between the visible (VIS) and near-infrared (NIR) bands, over their sum (Jensen, 2000). Written mathematically as

$$\text{NDVI} = (\text{NIR} - \text{VIS}) / (\text{NIR} + \text{VIS}) \quad (2.1)$$

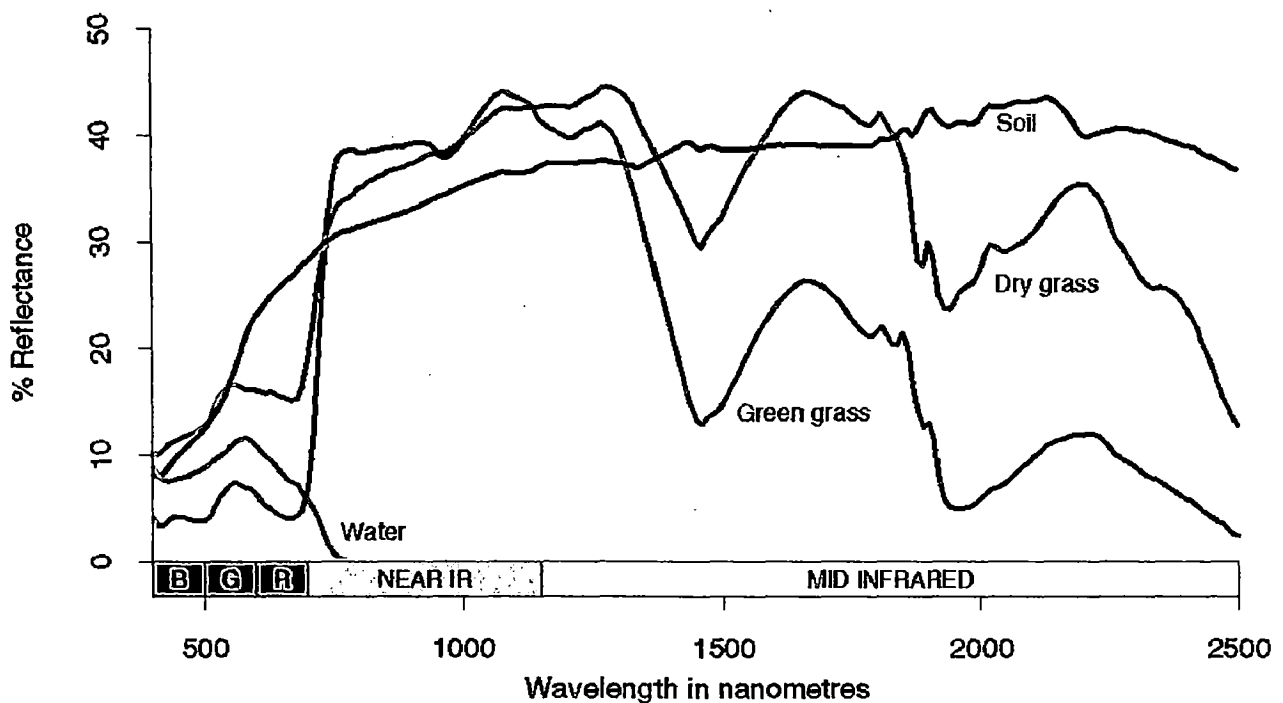


Fig 2.1 Reflectance characteristics of various objects with wavelength

Nearly all satellite Vegetation Indices employ this difference formula to quantify the density of plant growth on the Earth. The NDVI is an alternative measure of vegetation amount and

condition. It is associated with vegetation canopy characteristics such as biomass, leaf area index and percentage of vegetation cover.

When sunlight strikes objects, certain wavelengths of this spectrum are absorbed and other wavelengths are reflected. The pigment in plant leaves, chlorophyll, strongly absorbs visible light (from 0.4 to 0.7  $\mu\text{m}$ ) for use in photosynthesis. The cell structure of the leaves, on the other hand, strongly reflects near-infrared light (from 0.7 to 1.1  $\mu\text{m}$ ). The more leaves a plant has, the more these wavelengths of light are affected. Fig. 2.1 gives reflectance characteristics of various objects with wavelength. Vegetation appears very different at visible and near-infrared wavelengths. In visible light, vegetated areas are very dark, almost black, while desert regions (like the Sahara) are light (see Fig. 2.1). At near-infrared wavelengths, the vegetation is brighter and deserts are about the same. By comparing visible and infrared light, scientists measure the relative amount of vegetation. In general, if there is much more reflected radiation in near-infrared wavelengths than in visible wavelengths, then the vegetation in that pixel is likely to be dense and may contain some type of forest. If there is very little difference in the intensity of visible and near-infrared wavelengths reflected, then the vegetation is probably sparse and may consist of grassland, tundra, or desert.

Calculations of NDVI for a given pixel always result in a number that ranges from minus one (-1) to plus one (+1). A zero means no vegetation and close to +1 (0.8 - 0.9) indicates the highest possible density of green leaves. Normally for vegetation NDVI values typically range between 0.1 and 0.7. Higher index values are associated with higher levels of healthy vegetation cover, whereas clouds and snow will cause index values near zero, making it appear that the vegetation is less green.

### **2.3 Hydrologic Modeling**

*Renschler (1996) defines a model as a simplification of processes and their interactions with the aim of extracting, evaluating and simulating the relevant processes.* The different types of hydrologic response models were proposed, ranging from purely empirical simple models to highly sophisticated distributed physical process models. Based on the degree of representation of the physical processes, the models are classified with the increasing degree

of representation, as Empirical models, physically based lumped and distributed model and Neuromorphic models.

### **2.3.1 Empirical models**

They are based primarily on observation and seek to characterise system response from those data. An empirical model does not attempt in any way to represent the process occurring within the catchment, not even in a simplified manner. But in many situations in practice empirical models are used. However these simpler models normally fail to represent the non-linear dynamics, which are inherent in the process of rainfall-runoff transformation and sediment yield process.

#### ***(a) Rainfall-Runoff Models***

The origin of empirical hydraulic modeling dates back to the rational method developed by Mulvaney (1850) and an 'event' model by Imbeau (1890) for relating storm runoff peak to rainfall intensity. Other previous investigators who proposed empirical or semi-empirical relationships for runoff and rainfall are Sherman (1932), Horton (1930, 1939), Izart (1944), Horton (1945) and soil conservation services (SCS) (1956) (Singh, V.P., et. al 2002).

Khosla analysed the rainfall runoff and temperature data for various catchments in India and USA and derived an empirical equation for calculating monthly runoff. Based on analysis of data from small sub-catchments of different rivers basin in India (Subramanya, 1994), Kothyari (1995) proposed an empirical relationship for estimation of monthly runoff considering concurrent and antecedent monthly rainfall as inputs.

Jacob (1995) developed a regional regression equation for predicting monthly runoff which is applicable over the whole country of Nepal. While deriving this equation monthly data was used in a multiple regression analysis involving up to 14 catchment parameters such as basin area, main stream length, area of catchment below 5000 m in elevation etc. A set of 12 regression equations were derived which can be used to predict the mean flow in each month of the year. The general form of such equation is



$$Q_{\text{mean (month)}} (\text{m}^3/\text{s}) = \text{Coefficient } C \times (\text{Basin Area})^{a_1} \times (\text{Basin area below 5000m} + 1)^{a_2} \times (\text{Mean monsoon Precipitation})^{a_3} \quad (2.2)$$

Where  $a_1$ ,  $a_2$  and  $a_3$  are constant and are different for different months. Basin area is measured in square kilometers and the mean monsoon precipitation is given in millimeters.

Saranghi et al. (2005) have developed a regression model for St. Esprit watershed, Quebec, Canada using watershed-scale geomorphologic parameters to predict surface runoff and sediment losses. Morphological parameters such as bifurcation ratio  $R_B$ , area ratio  $R_A$ , channel length ratio  $R_L$ , drainage factor  $D_F$  and relief ratio  $R_R$  were selected using the Multivariate Adaptive Regression Splines tool, based on their relative importance in prediction of runoff  $Q$  (in  $\text{m}^3/\text{s}$ ). The form of equation they obtained is as follows:

$$Q = 0.13 R^{\sqrt{R_B}} - 0.87 R^{\sqrt{R_L}} - 0.02 R^{\sqrt{R_A}} + 4.631 R^{\sqrt{D_F}} - 48.15 R^{\sqrt{R_R}} + 45.5 \quad (2.3)$$

Where  $R$  is event based rainfall in mm.

### (b) *Sediment Yield Models*

Most of the models used in soil erosion studies are empirical models. The most widely used empirical model is the universal soil loss equation (USLE). Others include revised universal soil loss equation (RUSLE), modified universal soil loss equation (MUSLE) etc, which are based on modifications made on USLE. The USLE (Wishmeier and Smith, 1978) is the most widely used model in predicting soil erosion. Since the model was developed based on simulation in the East of the Rocky Mountains, its validity in areas outside the USA is regularly questioned (Roo, 1993). The USLE model estimates average annual soil loss by sheet and rill erosion on those portions of landscape profiles where erosion but not deposition is occurring. The model does neither predict single storm event nor does it predict gully erosion (Foster, 1982; Keneth et al., 1991). The model is also one-dimensional and static with limited possibilities for analysis of phenomenon dynamics (Jaroslav et al., 1996).

MUSLE is one of the modified versions of the USLE. In MUSLE, the rainfall energy factor was replaced with runoff. The runoff factor includes both total storm runoff volume and peak runoff rate. Compared to USLE, this model is applicable to individual storms, and eliminates the need for sediment delivery ratios, because the runoff factor represents energy used in detaching and transporting sediment. The main limitation is that it does not provide information on time distribution of sediment yield during a runoff event. It is strictly a sediment yield equation and should not be used where detachment controls sediment yield (Roo, 1993).

RUSLE is a revised version of USLE, intended to provide more accurate estimates of erosion (Renard et al., 1997). It contains the same factors as USLE, but all equations used to obtain factor values have been revised. It updates the content and incorporates new material that has been available informally or from scattered research reports and professional journals. The major revisions occur in the cover management factor (C), support practice factor (P), and slope length gradient factor (LS) factors. The C is now the product of four sub factors: prior land use, canopy cover, soil surface cover and surface roughness (Yazidhi, 2003).

The Morgan, Morgan and Finney (MMF) model is another empirical model for predicting annual soil loss from field-sized area on hill slopes (Morgan et al., 1984). The model separates the soil erosion process into two phases i.e. the water phase and the sediment phase. In the water phase annual rainfall is used to determine the energy of the rainfall for splash detachment and the volume of runoff, assuming that runoff occurs whenever the daily rainfall exceeds a critical value representing moisture storage capacity of the soil-crop complex and that the daily rainfall amounts approximate an exponential frequency distribution. In the sediment phase, splash detachment is modeled using a power relationship with rainfall energy modified to allow for the rainfall interception effect of the crop. The model has been revised with new changes incorporated owing to the rise in data availability and difficulties in estimating certain parameters as in the original version. In the revised version, changes have been made to the way soil particle detachment by raindrop impact is simulated, which now takes account of plant canopy height and leaf drainage, and a component has been added for soil particle detachment by flow. For the estimation of soil loss by this method following relation is used.

Soil loss = minimum value of the two: transport capacity of overland flow (G) and the estimated rate of soil detachment (F).

$$G = C * Q^2 * \sin S * 10^{-3} \quad (2.4)$$

$$F = K (E \exp^{-aP})^b . 10^{-3} \quad (2.5)$$

Where C, Q, S, K, E, and P are cover management factor, overland flow, sine of slope gradient, soil detachability index, kinetic energy of rainfall, percentage rainfall intercepted by crop respectively.

Shrestha(1997) has applied MMF approach in Likhu Khola valley of middle mountain in Nepal and could reasonably assess the soil loss from different categories of land use pattern.

### **2.3.2 Physically based model**

#### **(a) Lumped parameter conceptual model**

The lumped conceptual model occupies intermediate position between fully physically based distributed approach and empirical black box approach. In conceptual type model the internal description of the various sub process are modeled attempting to represent in a simplified way the known physical process. Even if not applying the exact differential laws of conservation, conceptual models attempts to describe large spatial and temporal scale conservation and response laws that are in accordance with the observed large scale behaviour of water in hydrologic drainage basin. Conceptual approaches improve the hydrological response of a basin in comparison to empirical equations. Conceptual models are developed assuming watershed to be fairly homogeneous, i.e. the spatial variability of input, transfer function and outputs are not explicitly taken into account. Because the inputs are averaged, or lumped, the model have come to be called “lumped” model.

#### **(b) Distributed Parameter Models**

The physically based distributed models are based on our understanding of the physics of the hydrological processes, which control the catchment response, and use physically based equations to describe these processes. Also, these models are spatially distributed since the

equations from which they are formed generally involve one or more space coordinates. A discretization of spatial and temporal coordinates is made and the solution is obtained at the nodal points of this discretized representation. This implies that these models can be used for forecasting the spatial as well temporal pattern of more than one hydrological variable. Such models require much of computational time and also require advance computers as well as a broad data base. In these models transfer of mass, momentum and energy are calculated directly from the governing partial differential equations which are solved using numerical methods, for example the St. Venant equations for surface flow, the Richards equation for unsaturated zone flow and the Boussinesq equation for ground water flow.

Though physically based hydrological models have proved to be very useful for simulation of various processes related to the management of water such as hydrodynamic, morphological, ecological, water quality, groundwater flow etc. implementation and calibration of such a models can typically present various difficulties (Duan et al, 1992), requiring sophisticated mathematical tools (Sorooshian et al, 1993), significant amounts of calibration data (Yapo et al, 1996), some degree of expertise and experience with models (Hsu et al, 1995).

Stanford Watershed Model –SWM (now HSPF) by Crawford and Linsley (1966) was probably the first physically based model to model virtually the entire Hydrologic cycle. Simultaneously a number of somewhat less comprehensive models were developed. Examples of such models that become popular are the watershed models of Dawly and O'Donnell (1965) and HEC-1 (Hydrologic Engineering Center 1968). Also number of semi distributed models capable of accounting for the spatial variability of hydrologic process within the watershed were developed as illustrated by Tank model developed by Sugawara (1967) (Singh, V.P., et. al 2002).

Indeed there has been a proliferation of watershed hydrology models since the development of Stanford watershed model (SWM), with emphasis of physically based models. Example of such Watershed hydrologic Models are storm water management model (SWMM) (Metcalf and eddy 1971), national weather service (NWS) river forecast system (Burnash 1973), streamflow synthesis and reservoir regulation (SSARR) (Rockwood 1982), System Hydrologique European (SHE) (Abbott 1986), TOPMODEL (Bevven and Kirkby 1979), institute of hydrology distributed model (IHDM) (Morris 1980) and so on .All of these models have since been significantly improved, SWM, now called hydrologic simulation

package-fortran (HSPF), is far more comprehensive than its original version. TOPMODEL has been extended to contain increased catchments information more physically based processes and improved parameter estimation ((Singh, V.P., et. al 2002).

Chemical, Runoff, and Erosion from Agricultural Management system (CREAMS ) was developed by U.S. department of agriculture (USDA) in 1980 fro simulating agricultural runoff and water quality. WEPP (Nearing et al., 1994) was developed for use in soil and water conservation and environmental planning and assessment. The Water Erosion Prediction Project (WEPP) erosion model computes estimates of net detachment and deposition using a steady state sediment continuity equation. Agriculture Non –Point Source Model (ANGPS) (Yang et. al 1989,1995) is an example of distributed parameter event based model which can be used to simulate water quality and quantity ((Singh, V.P., et. al 2002).

### **2.3.3 Nuromorphic models**

This is the soft computing approach of modeling the system .Soft computing referes to a consortium of computational methodologies. Some of its principal components include Arificial Neural Netwok (ANN), Fuzzy Logic (FL) and Genetic algorithm (GA), all having their roots in Arificial Intelligence (AI). Of these three components ANN and Fuzzy Logic approache are discussed in brief.

#### ***(a) Artificial Neural Network***

Artificial neural network (ANN) is a new soft computing technique composed of densely interconnected processing nodes, which has the ability to extract and store the information from the few patterns (data) in training through learning. ANN architecture parallels in processing with that designed to process the information in neuron computing . The model is easy to develop; yields satisfactory results when applied to complex systems poorly defined or implicitly understood; and more tolerant to variable, incomplete or ambiguous input data. Hydrologic applications of ANN include the modeling of rainfall-runoff forecasting, sediment yield process, snow-rainfall process, assessment of stream's ecological and

hydrological responses to climate change, and ground water quality prediction and ground water remediation.

Neural network (NN) plays big role in the field of water sector where complex natural processes dominate. The high degree of empiricism and approximation in the analysis of hydrologic response of watershed make the use of neural network highly suitable. In other words, when the possibility of representing the complex relationships between various aspects of the processes in terms of physical or conceptual modeling is very remote, the neural network plays an important role.

Previous works by Karunanithi et al. (1994), Dawson and Wilby (1998), Campolo et al. (1999), Zealand et al. (1999) and Imrie et al. (2000), have demonstrated the capability of ANN in streamflow forecasting. The ANNs, they used performed much better than the conventional models. The application of ANN for modeling the rainfall-runoff process started with preliminary work by Halff et al. (1993) who used a three layer feedforward ANN for the prediction of hydrograph. Since then, many studies in the context of rainfall-runoff modeling using ANNs have been carried out (Smith and Eli, 1995; Shamseldin, 1997; Tokar and Johnson, 1999). Similarly neural network (NN) approaches to model the streamflow-suspended sediment relationship were investigated by Kisi (2005) and Agrawal (2005). In many or most occasions it was shown that the neural networks tend to give better result than the deterministic models, provided that the process under consideration is not changed in time. The following paragraphs give the brief description of some of the sediment yield and Rainfall-Runoff ANN models.

Kisi (2005) evaluated the performance of ANN in estimating suspended sediment load at Rio Valenciano station operated by US Geological survey. He investigated four different scenarios with different combinations of runoff ( $Q$ ) and sediment discharge ( $Q_s$ ). The input combinations used in this application to estimate suspended sediment values for Rio Valenciano station were (i)  $Q_t$ ; (ii)  $Q_t$  and  $Q_{t-1}$ ; (iii)  $Q_t$  and  $Q_{st-1}$ ; and (iv)  $Q_t$ ,  $Q_{t-1}$  and  $Q_{st-1}$ . The node number in the hidden layer was found to vary between 1 and 2. Accordingly, an ANN structure ANN(3,2,1) consists of 3 inputs, 2 hidden and one output nodes. In this case the input layer covers the current and one antecedent flow and one antecedent sediment ( $Q_t$ ,  $Q_{t-1}$ ,  $Q_{st-1}$ ) and the output layer consists of the unique sediment concentration value at day  $t$ . The feedforward error back propagation algorithm was adopted for training the neural

network and the sigmoid and linear functions were used for the activation functions of the hidden and output nodes, respectively. On validation, the  $R^2$  value for above mentioned four scenarios were 0.867, 0.874, 0.702 and 0.876, respectively. This implies that ANN can perform well in forecasting suspended sediment yield.

Sajikumar and Thandaveswara (1998) applied the concept of ANN in rainfall-runoff modeling by utilizing the data of River Lee, UK and the Thuthapuzha River, India. Feed-forward multilayer perceptron (MLP) network with temporal back propagation algorithm was adopted in their study. The network architecture was designed consisting of single hidden layer with five neurons. Non linear sigmoid function was used as transfer function in all layers. The  $R^2$  value for this particular case was 0.799.

Jha and Jain (2005) also investigated use of ANNs in rainfall-runoff modeling in Kentucky River basin, USA. ANN model in this study consisted of three layers. The neurons input layer consisted of the total rainfall at times  $t$ ,  $t-1$ ,  $t-2$  and observed discharges at times  $t-1$ , and  $t-2$ . The only neuron in the output layer represented the flow at time  $t$ . The sigmoid activation function was used as the transfer function at both hidden and output layer. Standard back-propagation training algorithm with generalized delta rule was employed to train the neural networks. The result obtained in this study indicates that the single-hidden layer feed-forward ANN consisting of five input neurons, four hidden neurons, and one output neuron was best ANN model for the data set considered. The performance of the ANN rainfall-runoff models was excellent in terms of correlation coefficient, Nash-Sutcliffe efficiency and root mean square error (RMSE).

### ***(b) Fuzzy Logic System***

There are having lot of uncertain information, which are used to estimate and understand many complex problems. But to represent them mathematically is very difficult because computers or processors cannot understand what one perceive. To make these devices compatible with ones thinking, one needs to represent these uncertain information mathematically. For this purpose, to handle the concept of "partial truth", Dr. Lofti Zadeh in 1965 proposed a new theory called "Fuzzy Sets" (Zadeh, 1965). The logical processing using fuzzy sets, is known as "Fuzzy Logic"..

Many researchers have applied the fuzzy approach to various engineering problems (Mamdani 1974; Pappis and Mamdani 1977; Sugeno 1985; Sen 1998; Sen 2001). The basis of fuzzy logic is to consider hydrological variables in a linguistically uncertain manner, in the forms of subgroups, each of which is labeled with successive fuzzy word attachments such as “low”, “medium”, “high”, etc. In this way, the variable is considered not as a global and numerical quantity but in partial groups which provide better room for the justification of sub-relationships between two or more variables on the basis of fuzzy words.

Sen and Altunkaynak (2003) endeavoured to evaluate the performance of Fuzzy Logic approach in rainfall-runoff modeling for two different drainage basins on the European and Asian sides of Istanbul. In that study it was attempted to compare regression and fuzzy logic approach. It was investigated that fuzzy logic model prediction yields less relative error as compared to the regression method.



## CHAPTER –3

### DESCRIPTION OF THE STUDY AREA

---

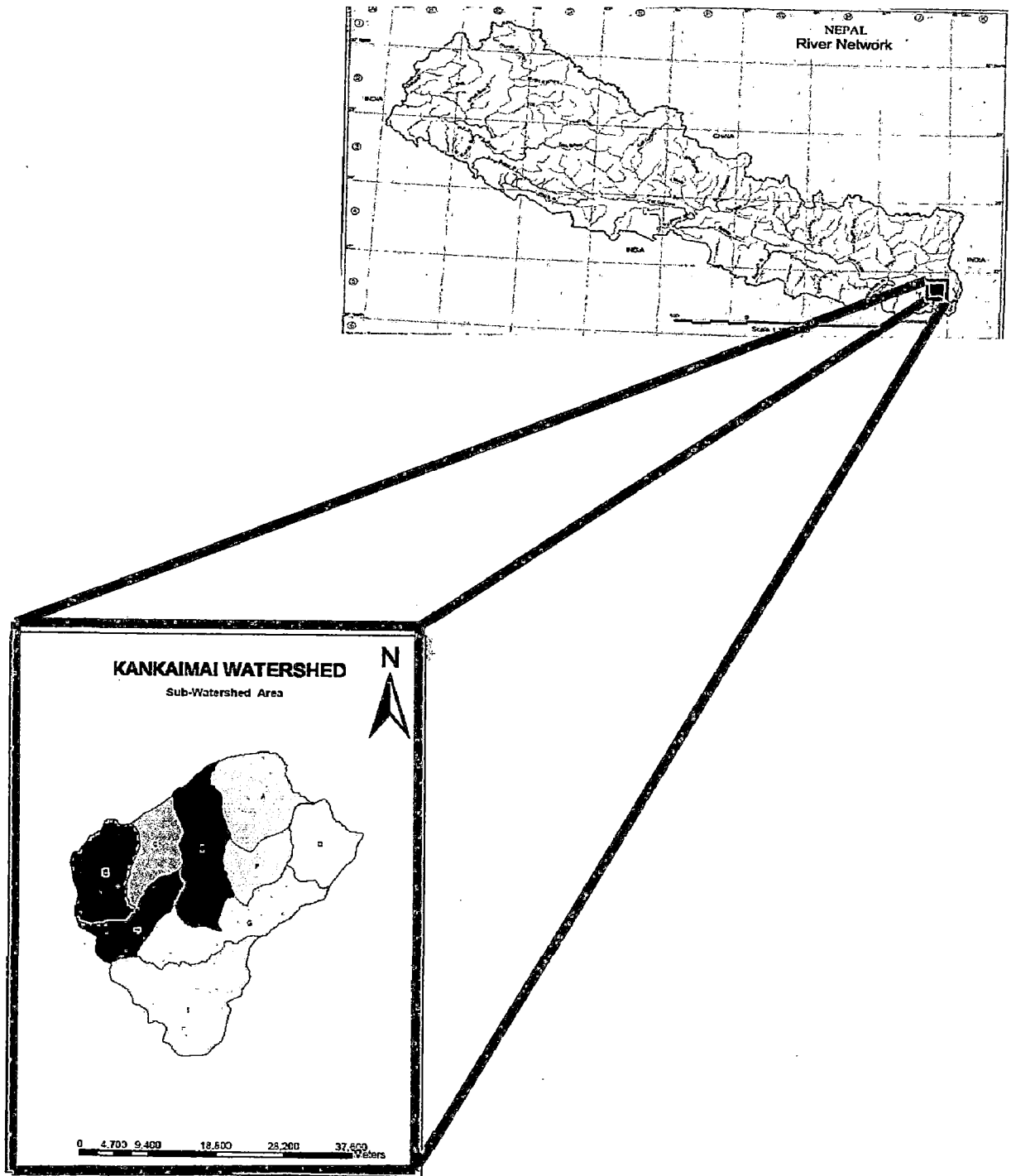
#### 3.1 General

The Kankaimai watershed, located in eastern Nepal, has been selected for the present study. Kankaimai river, the main river of basin, is a spring fed perennial river that originates from the Mahabharat range with name Mai Khola at an elevation of about 3300 m above mean sea level (msl). The highest altitude of the basin is 3,636 m above mean sea level (msl) where a village called Sattapur is located. The river flows north to south and takes the name of Kankaimai downstream of its confluence with Deumai Khola. From this reach onwards, the nature of the river changes from "confined bed rock type" to alluvial with flood plains. The Kankaimai becomes Kankai as it enters Jhapa District and downstream from this reach it is completely alluvial in nature with significant flood plains.

For the purpose of this study, the catchment area upstream of the Kankai Mai River at Mainachuli has been considered. The entire catchment area of 1180 km<sup>2</sup> lies within the district of Ilam. In terms of latitudes and longitudes, the overall catchment boundaries are 87°35' to 88°10' (latitude) and 26°37' to 27°05' (longitude).

The major tributaries of Kankaimai are Mai, Lodhiya Khola, Deumai Khola, Puwa Khola and Jogmai Khola. The total length of the river from Mainachuli to the source is about 90 km.

The Mai Khola meets Deumai Khola at Laramba, about 23 km upstream of Chepti. In this stretch, the river is relatively flatter having average gradient of about 1:235 compared to steep gradient in the upstream reach.



**Fig 3.1 Location Map of Study Area**

STUDY AREA

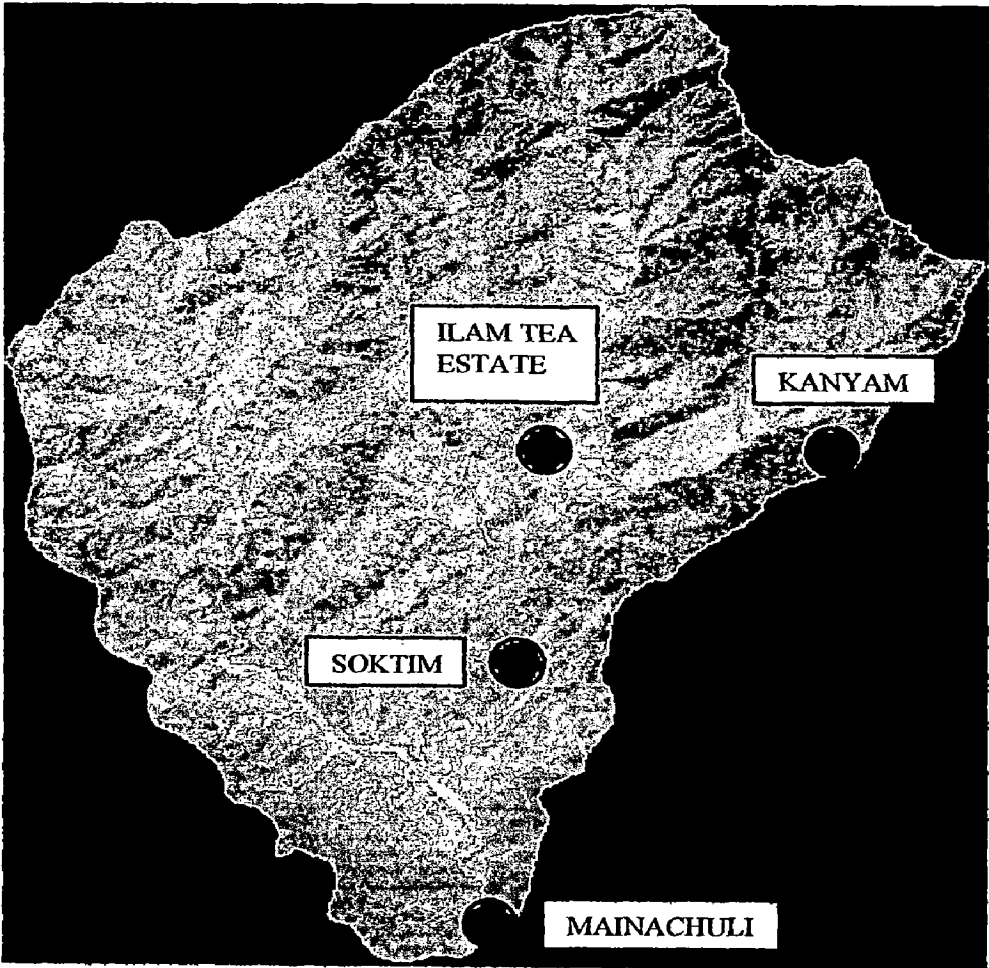


Fig. 3.2

### 3.2 River Characteristics

The confluence of Puwa *Khola* and Mai *Khola* is about 49 km upstream of Chepti (i.e., 26 km upstream from the confluence of Mai *Khola* and Deumai *Khola*). The river gradient between the confluence of Puwa *Khola* and Mai *Khola* to Deumai *Khola* and Mai *Khola* is about 1:75. The confluence of Jogmai *Khola* and Mai *Khola* is about 9 km upstream from the Puwa *Khola* and Mai *Khola* confluence (i.e., 58 km upstream from Chepti) and the river gradient along this stretch is about 1:75. Up to 30 km upstream of the confluence of Jogmai *Khola* and Mai *Khola*, the river gradient is steeper than 1:23 and the remaining stretch lies in the Mahabharat range.

### 3.3 Hydrological and Meteorological Data

#### (a) Climate

The catchment of Mai *Khola* lies in the Lesser Himalayas and Sub-Himalayas regions. The physiographic characteristics of Mai *Khola* and especially the variation in altitude influence the climate of the basin. Therefore, the climate of the catchment area varies from subtropical climate in the lower region to temperate climate in the upper region. As in other parts of Nepal, the catchment area of the river experiences southeast monsoon, which on average lasts from June to the end of the September. The average daily temperature recorded at the Ilam Tea State station, Kanyam Station and Saktim Station are taken in the present study. Mean monthly evaporation data, as well as mean temperatures, relative humidity and sunshine hours of Kankaimai basin are presented in Table 3.1. The average daily temperatures (1995-1999) recorded at Ilam Tea State, Saktim and Kanyam are given in Appendix A

**Table 3.1: Mean Monthly Temperature, Relative Humidity and Sunshine Hours of Kankaimai Basin**

Basin / Sub-Basin	Months											
	Jan	Feb	Mar	Apr	May	Jun	Jul.	Aug	Sept	Oct	Nov	Dec
Mean Temperature ( $^{\circ}$ C)	12.6	14.0	18.3	20.7	21.2	22.2	22.1	22.5	21.7	20.2	17.1	14.1
Relative Humidity (%)	77	73	67	71	82	90	92	91	90	84	76	74
Sunshine Hours (hrs)	7.2	6.4	6.3	7.6	5.9	2.9	1.3	2.7	2.9	6.7	6.7	8.7
Evaporation (mm/day)	2.86	3.94	5.92	6.58	5.96	5.20	4.20	4.39	4.11	4.27	3.31	2.92

## (b) Rainfall

The rainfall station at Ilam Tea State, Kanyam and Saktim established by Department of Hydrology and Meteorology (DHM) Nepal which are located in the basin and its periphery are selected for runoff analysis. The mean monthly rainfall at different stations in and around Kankaimai basin are given in table 3.2 and daily rainfall recorded at three stations are given in Appendix B.

**Table 3.2: Mean Monthly Rainfall of Stations in and around Kankaimai Basin (mm)**

Station Index number	Location	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Total
1406	Memeng Ghat	18.7	20.9	48.5	120.9	226.4	331.8	481.9	406.8	282.4	105.9	17.5	12.5	2074.2
1407	Ilam Tea Estate	12.4	13	22.6	60.8	136.9	299.8	432	318.3	233.1	66.5	11.1	6.7	1613.2
1408	Damak	13.4	13.5	22	64.4	182.6	431.8	706.8	555.6	357.5	131.4	15.5	6.3	2500.8
1410	Himali Gaon	14.7	19.7	30.1	80.7	166.5	429.6	638.5	450.2	347.7	86.1	18.5	12.4	2294.7
1415	Sanischar e	12.4	17.8	25.7	66.1	210.9	530.8	803.1	550.5	415	114.7	22.1	8.5	2777.6
1419	Phidim	13.9	17.5	34.3	77.8	150.2	179.4	346.8	249.4	152	36.9	13.4	14.6	1286.2
	Average	14.3	17.1	30.5	78.5	178.9	367.2	568.2	421.8	298	90.3	16.4	10.2	2091.1

As can be seen from the above data, significant portion of the total rainfall in the catchment area occurs during the four months of monsoon, i.e. June to September. In fact, the monsoon contributes almost 79 percent of the total annual rainfall. On the other hand, only about 3 percent of the total annual rainfall occurs during the dry / lean season (November to February). The entire hydro-meteorological characteristics of the basin are strongly characterised by the high precipitation generating peak monsoon flows and low precipitation during the dry season resulting in low flows. Water availability, occurrences of extreme events - both low and high and other related aspects in the river basin are governed by these factors to a large extent.

### (c) Runoff

DHM has installed three stream gauging stations (hydrometric station) on the Kankaimai basin. The details of these stations are reported in Table 3.3.

**Table 3.3: Stream Gauge Stations at and around the Kankaimai Basin**

Station Index No.	River Name and Location	Elevation (m)	Lat./Long.	Years of Records	Drainage Area (km <sup>2</sup> )
728	Mai Khola at Rajduwali	609	26°52'45"/87°55'45"	1983-1999	377
730	Puwa Khola at Sajbote	802	26°55'00"/87°54'40"	1965-2000	107
795	Kankai Mai River at Mainachuli	125	26°41'12"/87°52'44"	1971-1999	1180

Hydrologic response studies i.e. water and sediment yield prediction of Kankaimai basin is based on the gauging station No. 795 at Mainachuli. The daily runoff data from the year 1995 to 1999 have been used for daily runoff prediction and that from 2001 to 2003 has been used for the sediment yield prediction.

The summaries of long-term mean monthly runoff of Kankaimai basin is given in the Table 3.4 and daily runoff data at mainachuli in the year 1995 to 1999 has been presented in Appendix C.

**Table 3.4: Mean Monthly Runoff of Kankaimai Basin and Sub-basins**

Basin and Sub-Basin	Index No. (Station)	Month /Discharge (m <sup>3</sup> /s)											
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mai Khola	795	12.10	9.90	8.86	10.74	19.43	62.96	179.24	165.51	129.73	55.51	26.45	17.67
Puwa Khola	730	2.84	2.44	1.98	2.20	3.38	7.04	12.72	13.22	11.96	9.21	5.67	3.84
Mai Khola	728	5.89	4.46	4.48	6.27	10.47	31.92	63.87	71.51	55.47	22.53	8.87	5.64

#### (d) Sediment Yield Data

The department of hydrology and meteorology has recently established a gauging station for the measurement of suspended sediment in Kankaimai river at Mainachuli. At present data only in wet season from 2001 to 2003 are available and same has been adopted for the simulation purpose. The average monthly suspended sediment yield is shown in table 3.5 and daily data alongwith corresponding daily runoff is presented in Appendix D.

**Table 3.5 Monthly Suspended Sediment Yield (t/day)**

Months/Year	July	Aug	Sep	Oct	Nov	Dec
2001	4000	6166	6021	17837	167	52
2002	127189	11151	1723	-	-	-
2003	39831	10361	3307	3813	-	-

### 3.4 Land Resources

In terms of land resources, the Kankaimai basin is covered with forest, cultivated land, tea gardens, settlements, water bodies, grazing land, sand bars, barren land and swampy areas. These are broadly categorised into five groups namely; forest land, cultivated land, grazing land, shrub land and others. Details of this broadly classified land utilisation are discussed in the following subsections.

**Forest land:** Forests in the Kankaimai basin consist of coniferous, hardwood and several other combinations of tree species. In addition, all range of crown densities of trees including protected and non-protected forests exist. The forest cover of the Kankaimai basin is about 37% of total basin area considered in this study.

**Cultivated land:** Cultivated land in the basin consists of Hill-slope cultivation, Level-terrace cultivation and Valley cultivation. In valley cultivation land type, cultivation valley floor, *tars* (*large flat areas along river banks*), foot-slopes and alluvial fans are incorporated. In addition, tea gardens are also included in the slope cultivation category. The percentage of cultivated land area is about 35% of the total Kankaimai basin considered in this study.

**Grazing land:** Grazing lands exist in all temperate zones, namely; Sub-tropical, warm temperate, temperate and cool temperate zones. 2% of the total Kankaimai basin considered in this study is covered by grazing land.

**Shrubs:** Shrub land includes shrubs and shrub vegetation as well as hardwood regeneration. The percentage of shrub area is 13% of the total Kankaimai basin considered in this study.

**Other land types:** Settlement areas (including rural and urban), barren land, land covered by water bodies, sandbars and other wasteland are classified under 'other land'. The total area covered by the Other Land category is about 14,508 hectare, which is about 13% of the total basin area of Mai Khola. Within the Other Land, the barren land comprises about 2.5% and settlement area comprises less than 0.1% of total basin area.

### **3.5 Existing Water Resources Utilization**

The Water Resources of the Kankaimai basin are currently being used mainly for three purposes, namely irrigation, power generation and piped water (i.e., drinking water) supply.

**Irrigation:** Irrigation is the largest water resources utilisation sector in the Ilam District. According to updated inventories based on Mechi Hill Project (MHP), Divisional Office of Department of Irrigation (DOI) at Ilam/Jhapa and Water Inventory Report of Water Energy Commission Secretariate (WECS), there are about 489 irrigation systems within Ilam District. Out of these only 24 are implemented under Irrigation Sector Project (ISP) and Second Irrigation Sector Project (SISP) programs and the rest are Farmers Managed Irrigation System (FMIS). Khola Basin There are about 364 irrigation systems that use the surface water from Kankaimai and its tributaries. Out of these, 348 systems are farmer managed and 16 are government-assisted systems implemented under various programs.

**Drinking water:** Apart from irrigation, drinking water is largest consumptive use of the water resources in the basin. A significant population of the Ilam municipality has access to drinking water supply. The municipality abstracts water from Bhadi Khola and Bhalu Base spring sources for all its drinking water needs. Furthermore, there are 38 Village Development Committes.

(VDCs) in the Kankaimai basin that have access to drinking water systems. The water resources used are mostly small springs and the systems are mostly small and scattered.



**Hydropower:** The largest hydropower plant in Kankaimai basin is the Nepal Electricity Authority (NEA) owned 6 MW Puwa Small Hydropower Plant Besides the NEA Puwa Khola Small Hydropower Plant, 15 micro-hydro plants have been constructed in the Ilam district, out of which 14 are in the Kankaimai basin. In addition to the micro-hydropower plants, there are numerous Peltric sets in Ilam District. These are hydropower plants with installed capacity limited to 3 kW.

**Other usage:** Water mills ("*Ghatta*") are one of the indigenous technologies used in Nepal for grinding grains such as maize, millet and wheat.. In Kankaimai basin, there are about 12 water mills. As reported by Namsaling Community Development Centre (NCDC), all of these 12 water mills divert nominal flows (30-40 lps).

Many of the ceremonies of both Hindu and Buddhist religions are performed on the banks of rivers. Besides, people also bathe in the river during certain auspicious days of the year for religious purpose. The devotees offer prayers after a holy dip in the river. For these activities, substantial river discharge is needed. Besides utilisation of water resources for religious purpose, other usages are livestock watering and fishing..

Overall, not a single basin transfer case was reported for micro hydropower schemes, Peltric Sets and water mills. The only significant inter basin water transfer case is due to the Puwa Khola Plant.

## **CHAPTER-4**

# **GEOMORPHOLOGIC & VEGETATIONAL ANALYSES**

### **4.1 Background**

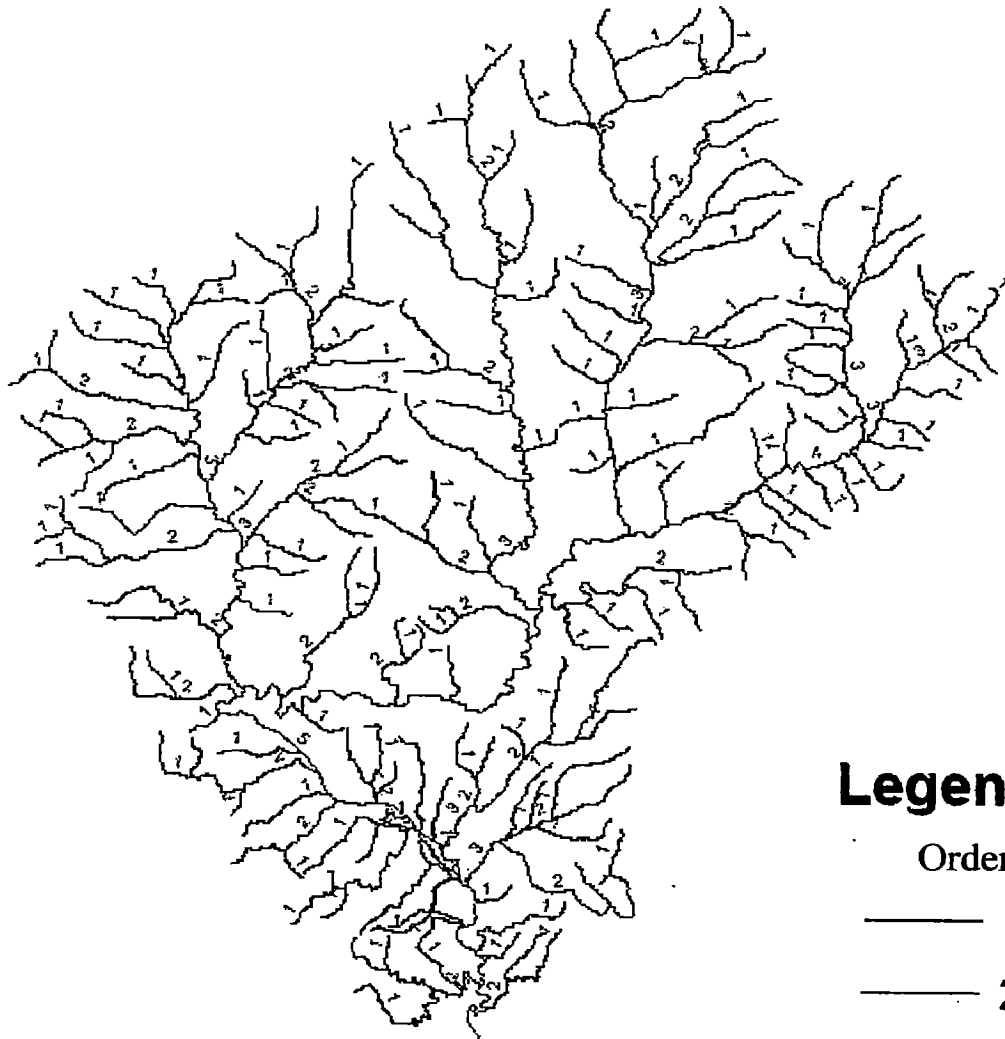
Morphological characteristics like stream order, drainage density, watershed length and width, channel length, channel slope and relief aspects of watershed are important in understanding the hydrological response of the watershed. Runoff response of the watershed is different for different slopes, shapes, lengths, widths and areas of watershed. Response is also affected by the factors like drainage density, length of overland flow, stream frequency, relative relief and relief ratios. Computation of watershed morphological characteristics is prerequisite to further detailed hydrological analysis of the watershed. Hydrologists have attempted to relate the hydrologic response of watersheds to watershed morphologic characteristics.

Similarly status of vegetation affects the response of watershed. Erosion and runoff reduces with increase in vegetation density. If the vegetation on the watershed is sparse, then there will be more chances of developing drainage channel thereby increasing drainage density and bifurcation ratio which have direct impact on the response of watershed.

### **4.2 Geomorphologic Analysis**

In the present study Geographical Information System (GIS) has been used for the computation of morphological characteristics of the watershed. Topographical map of 1:250000 scales were used as a base map for extracting various parameter required for the morphological analysis. The map was first scanned and georeferenced in UTM map projection system using four Ground Control Points (GCP). The ERDAS Imagine 8.3.1 software was used for registration purpose .Then on-screen digitization of watershed boundary, drainage network and contour line was carried out using ARC GIS 8.3 software.

# KANKAIMAI WATERSHED



## Legend

Order

- 1
- 2
- 3
- 4
- 5

**FIG 4.1 STRAHLER RIVER ORDERING SYSTEM**

Before scanning of topographical map watershed area was delineated with pencil with the help of contour lines and same was digitized to extract drainage basin layer. The Kankaimai watershed has fifth order stream system. The Horton-Strahler classification system was used for ordering of streams. The digitization work was started from the lower order stream and was sequentially proceed towards higher order streams. Fig. 4.1 shows the drainage network of the basin indicating stream of different ordering system. Finally data base was created for the different layers by adding respective attributes. Then by using clean and build command different geometrical parameter like length of streams of different orders, perimeter and area of drainage basin were computed.

**(a) Stream Lengths and Orders**

The most commonly used classification of stream network is the Horton –Strahler classification system. Accordingly, a first order stream is the smallest unbranched stream. The first order stream originates at the source. Two first order stream joins to form a second order stream. Likewise when two streams of order  $w_s$  join, a stream of order  $w_s+1$  is created. On the other hand, when two streams of different orders, say  $w_s$  and  $w_s+1$ , join the stream segment immediately downstream retains the higher of the orders of the two combining streams and consequently have the order  $w_s+1$  (Singh,V.P., 1989).

The length of streams of different orders which was calculated using GIS tool is summarized in Table 4.1. The Table also includes mean length of each river order streams and their respective percentage to the total stream length. Using these data following relationships were established (Fig 4.2.) between stream orders,  $w$  and number of stream,  $N$  and between stream order,  $w$  and length of streams,  $L$ , (Fig 4.3) respectively.

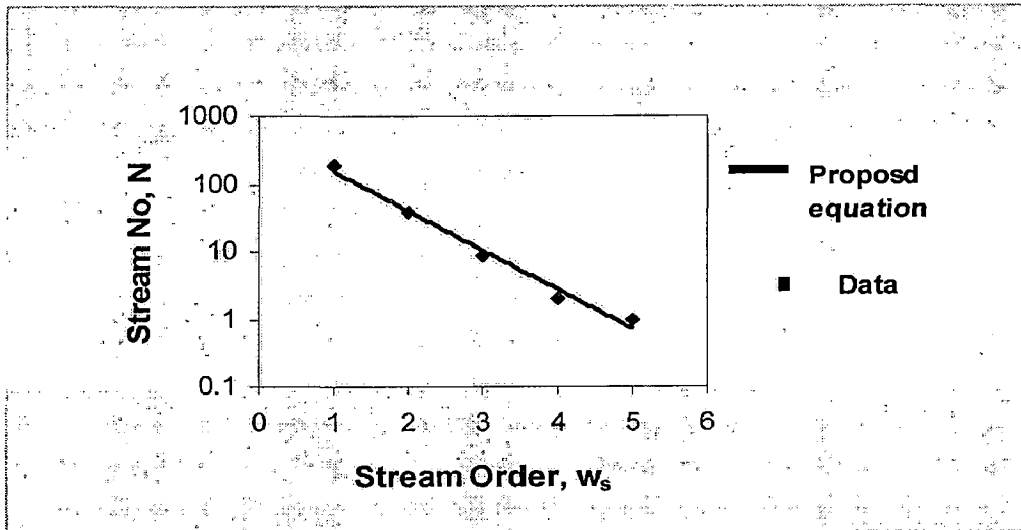
$$N = 581e^{-1.34w_s}; R^2 = 0.985 \tag{4.1}$$

$$L = 8.8 \times 10^5 e^{-0.71w_s}; R^2 = 0.960 \tag{4.2}$$

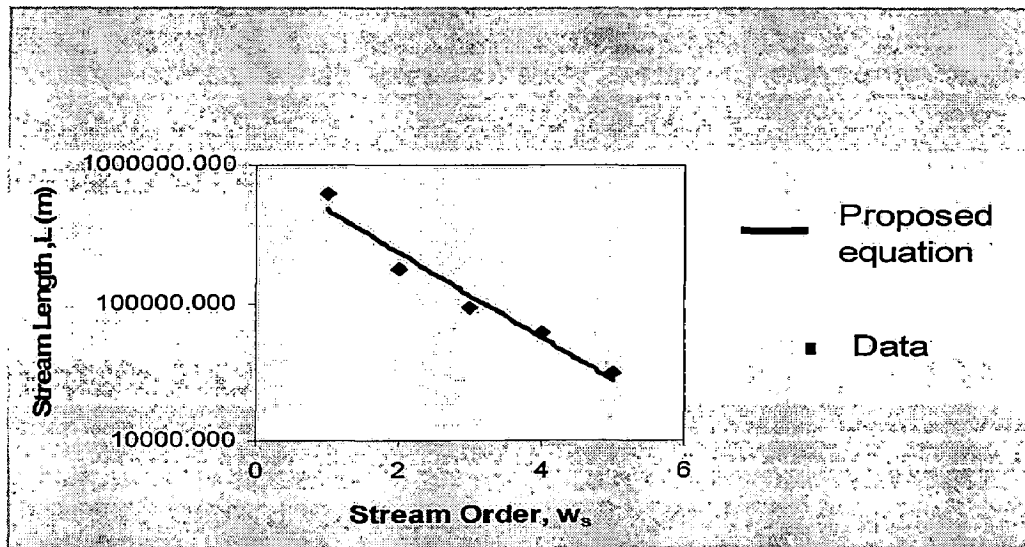
The above relationships are also shown through Fig. 4.2 and Fig. 4.3.

**Table 4.1 Details of Streams of Different Orders and Their Length**

Stream Order	No. of Stream Segment	Stream Length (m)	Mean Stream Length (m)	Percent to Total Stream Length
1	185	623348	3369	63.45
2	38	176637	4648	17.98
3	9	91844	10204	9.35
4	2	60166	30083	6.12
5	1	30424	30424	3.10
<b>Total</b>	<b>235</b>	<b>982422</b>		<b>100</b>



**Fig. 4.2 Relationship between Stream Order and Stream Number**



**Fig. 4.3 Relationship between Stream Order and Stream Length**

**(b) Bifurcation Ratio**

It is the ratio of the streams of lower order to the number of streams of the next higher order. This is a dimensionless quantity and shows only a small variation from one region to another. Bifurcation ratio is normally between 3 and 5 for the watershed where the influence of geological structure on drainage network is negligible. High bifurcation ratios are common where the effect of geological structure is dominant, but the shape of the basins also has an important effect. Abnormal bifurcation ratio usually has a marked effect on maximum flood discharge. Similarly higher bifurcation ratio between first and second order stream is the indicative for a state of accelerated erosion. If ground water table is deep and infiltration is high, surface runoff will be relatively low and less channel and a lower bifurcation ratio.

$$\text{Bifurcation Ratio, } R_b = N_{w_s} / N_{(w_s+1)}$$

(4.3)

Where  $N_{w_s}$  is the number of streams of  $w_s^{\text{th}}$  order and  $N_{w_s}$  is the number of streams of  $(n+1)^{\text{th}}$  order .

First order to second order,  $R_b = N_1 / N_2 = 185/38 = 4.87$

Second order to third order,  $R_b = N_2 / N_3 = 38/9 = 4.22$

Third order to fourth order,  $R_b = N_3 / N_4 = 9/2 = 4.5$

Fourth order to fifth order,  $R_b = N_4 / N_5 = 2/1 = 2.0$

Average bifurcation ratio = 3.9

**(c) Drainage Density**

Drainage density, which is defined as total stream length per unit area represents the degree of fluvial dissection. It is a valuable indicator of the relation between climate, vegetation, and the resistance of rock and soil to erosion. Under similar climate conditions, impervious rocks support a higher drainage density compared with permeable rock. Semi-arid areas have higher drainage densities than arid and humid areas with the same geology because of the rapid runoff and sparse vegetation. It is an index of the relative proportion of overland and channel flow. It also exerts a strong influence upon sediment yield and runoff response.

Drainage density can be expressed as:

$$D_d = \text{Total length of all streams} / \text{Total watershed area}$$

$$D_d = \sum_{i=1}^w \sum_{j=1}^{N_i} L_{ij} / A_w \tag{4.4}$$

$$D_d = 982.421 / 1180.00 = .832 \text{ km}^{-1}$$

**(d) Constant of Channel Maintenance**

As per the Schumm (1956) constant of channel maintenance can be expressed as the ratio of the drainage basin area to the total length of streams .

Mathematically,

$$C_c = A_w / \sum_{i=1}^{w_i} \sum_{j=1}^{N_i} L_{ij} \quad (4.5)$$

$$C_c = 1180/982.421 = 1.20$$

This implies that to maintain one km of stream 1.2 km<sup>2</sup> land area is required.

#### (e) Elongation Ratio

It is defined as the diameter of a circle with the same area as that of the basin to the maximum basin length. The elongation ratio is equal to one for a circular basin and approaches zero for a straight line. The more elongated basin has a dampened response, compared to the 'flashier' hydrograph of the circular shaped basin. This pattern results because of the first basin has a much broader distribution of flow path lengths, and therefore, a wide range of travel times.

Elongation ratio can be mathematically expressed as :

$$R_e = D/L_b = 38772/49817 = 0.778 \quad (4.6)$$

Diameter of circle, D having area of 1180 km<sup>2</sup> = 38772 m

Longest length of basin, L<sub>b</sub> = 49817 m

#### (f) Circularity Ratio

It is the ratio of drainage area A<sub>w</sub> and area of a circle A<sub>c</sub>, possessing the same perimeter of the basin, L<sub>p</sub>, i.e.,

$$R_c = A_w/A_c = A_w * 4\pi / L_p^2 = 1180 * 4\pi / 153.118^2 = .63 \quad (4.7)$$

#### (g) Channel segment frequency

Horton (1945) defined the channel -segment or stream frequency as the number of streams per unit area. This parameter is useful in determining hydrologic response of a basin. It can determine the length of overland flow and, in turn, the time of concentration.

Mathematically, it can be expressed as



$$C_f = \sum_{i=1}^w N_i / A_w = 235 / 1180 = .199 \text{ km}^2 \quad (4.8)$$

**(h) The Relief Ratio:**

It can be defined as the maximum relief divided by the longest flow-path, It indicates the overall steepness of a basin and can be related to the hydrologic characteristics.

$$R_h = ( H_{\max} - H_{\min} ) / L_b = (3636-125)/49817 = 0 .0705 \quad (4.9)$$

Where  $H_{\max}$  and  $H_{\min}$  are the maximum elevation of most distant ridge and elevation of basin outlet respectively.

**(i) The Relative Relief**

Melton defined relative relief as the ratio of basin relief to the length of perimeter .It is an indicator of general steepness of a basin from summit to mouth. It has an advantage over relief ratio in that it does not depend on basin length, which is questionable parameter in oddly shaped basin.

It can also be written in mathematical form as :

$$R_p = ( H_{\max} - H_{\min} ) / L_p = (3636-125)/153118 = .023 \quad (4.10)$$

**(j) Form Factor**

It is defined as the ratio of basin area to the square of basin length. Form factor is influenced by basin shape.

Mathematically,

$$R_f = A_w / L_b^2 = 1180/49.187^2 = .487 \quad (4.11)$$

### (k) Time of Concentration

This is the time it takes for water to travel from the most distance point of a watershed to the watershed outlet.

Using Kirpich formula (Subramanya, K., 1994).

$$\begin{aligned} T_c &= 0.01947(L_c^{0.77} / S_w^{.385}) && (4.12) \\ &= 0.01947 * \{90000^{0.77} / (3511/90000)^{.385}\} \\ &= 443 \text{ minutes} \\ &= 7.38 \text{ hours} \end{aligned}$$

Where

$L_c$  = Length of drainage line

$S_w$  = Slope of watershed

### 4.3 Vegetational Analysis

The Normalized Difference Vegetation Index (NDVI) has been used to categorize the different types of vegetation on the watershed. The georeferenced LANDSAT TM satellite image of 30 m resolution was acquired from the forestry department, Nepal government, for this analysis. The study area of interest was masked with the help of watershed boundary which has been delineated as discussed in section 4.2. The complete NDVI analysis was carried in ERDAS Imagine 8.6 (image processing software) environment.

In LANDSAT TM image, NIR is represented by band 4 and VIS is represented by band 3. Hence from Equation (2.1),

$$NDVI = (BAND 4 - BAND3)/(BAND 4 + BAND3) \quad (4.13)$$

NDVI values have been assigned for the different types of land covers as shown in Table 4.2

# LAND COVER MAP

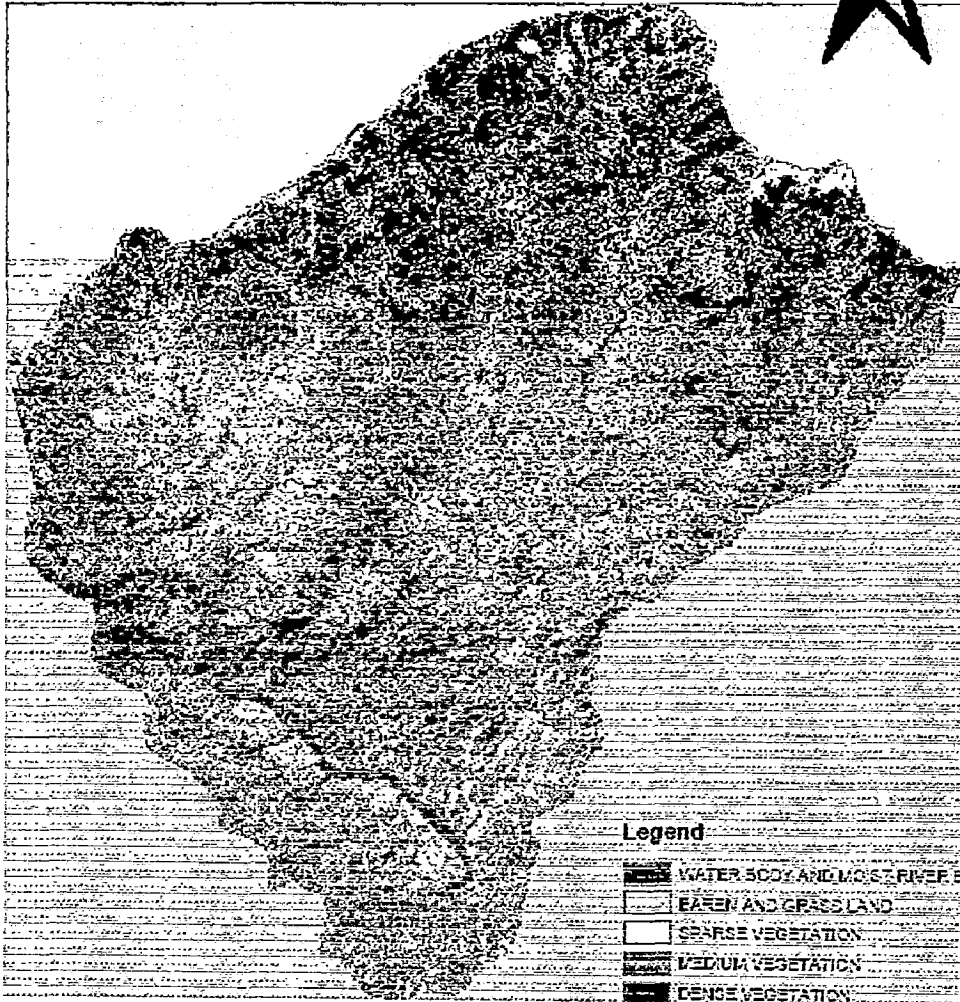


Fig. 4.4

**Table 4.2 Land Coverage and NDVI Values**

S.N.	Land Covers	NDVI Values
1	Water Body and moist river bed	-0.0058 to - 1
2	Barren land and grass covered area	0 .00093 to 0.123
3	Sparse vegetation	0.124 to 0.30
4	Medium vegetation	0.31 to 0.45
5	Dense Vegetation	0.46 to 0.74

The percentage coverage of above mentioned land covers classes is presented in table 4.3 and Fig 4.4 shows distribution of different categories of vegetation and other land cover types within the watershed area.

**Table 4.3 Land Coverage in Percentage**

S.N.	Land Cover Type	Percentage(Coverage)
1	Water Body and moist river bed	5.0
2	Barren land and grass covered area	9.0
3	Sparse vegetation	34
4	Medium vegetation	36
5	Dense vegetation	16

## CHAPTER – 5

# ANN MODEL DEVELOPMENT

---

### 5.1 Introduction of Artificial Neural Network

#### 5.1.1 General

An Artificial Neural Network (ANN) is an information processing paradigm which is inspired by the way biological nervous system, such as the brain process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of large number of highly interconnected processing elements (neurons) working in union to solve specific problems. An ANN is configured for specific application, such as pattern recognition or data classification, through learning process. Learning in biological systems involves adjustments to the synaptic connection that exist between the neurons. ANNs too imitate the way a human brain works.

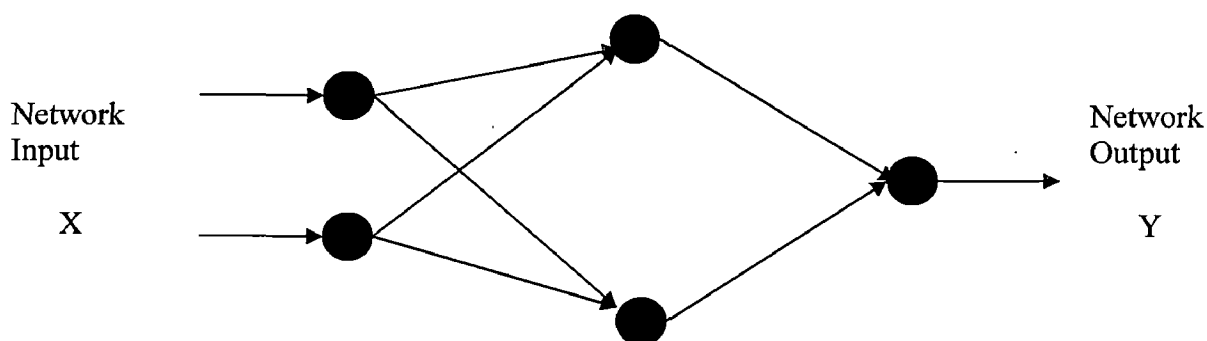
A neural network is a massively parallel –distributed processor that has a natural propensity or storing experimental knowledge and making it available or use .It resembles the brain in two respects (Sajikumar et al. 1999):

- (a) Knowledge is acquired by the network through a learning process and
- (b) Inter-neuron connection strength known as synaptic weight are used to store the knowledge.

ANNs have been developed as generalization of mathematical models of human cognition or neural biology. Their development is based on the following (Govindaraju et al. 1999).

1. Information processing occurs at many single elements called nodes, also referred as units, cells or neurons
2. Signals are passed between nodes through connection links.
3. Each connection link has as associated weight that represents its connection strength.
4. Each node typically applies a linear or non linear transformation called activation function to its net input to determine its output signal.

A neural network is characterized by its architecture that represents the pattern of connection between nodes, its methods of determining the connection weights and the activation function (Fausett, 1994). A typical ANN consists of a number of nodes that are organized according to a particular arrangement. One way of classifying neural networks is by the number of layers: single (Hopfield nets), bi-layer (Carpenter /Grossberg Adaptive Resonance works), and multi-layer (back propagation network). ANNs can also be categorized based on the direction of information flow and processing. In a feed forward network, nodes are generally arranged in layers, starting from first input layer and ending at the final output layer. There can be several hidden layers, with each layer having one or more nodes. Information passes from the input to the output side. The node in one layer is connected to those in the next, but not to those in the same layer. Thus, the output of a node in a layer is only dependent on the inputs it receives from previous layers and the corresponding weights. On the other hand, in a recurrent ANN, information flows through the nodes in both directions, from input to the output side and vice-versa. This is generally achieved by recycling previous network outputs as current inputs, thus allowing for feedback. Sometimes, lateral connections are used where nodes within layer are also connected. In the present study, feed forward networks were used. In most networks, the input layer receives the input variables for the problem at hand. This consists of all quantities that can influence the output .The input layer is thus transparent and is a means of providing information to the network. The last or output layer consists of values predicted by the network, and thus represents model output. The number of hidden layers and the number of nodes in each hidden layer are usually determined by a trial and error procedure. The nodes within neighbouring layers of the network are fully connected by links. A synaptic weight is assigned to each link to represent the relative connection strength of two nodes at both ends in predicting the input – output relationship. Figure 5.1 shows the configuration of a feed forward three layer ANN. In this figure,  $X$  is a system input vector composed of a number of causal variables that influence system behaviour,  $Y$  is the output vector composed of a number of resulting variables that represent the system behaviours.



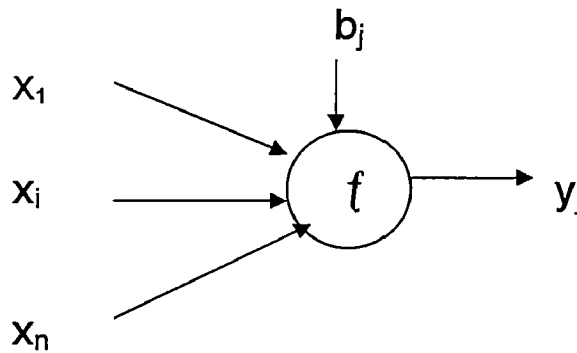
**Fig. 5.1 A Typical Configuration of feed forward three layers ANN**

### 5.1.2 Mathematical Aspects

A schematic diagram of a typical  $j^{\text{th}}$  node is displayed in Figure 5.2. The input to such a node may come from system causal variables or outputs of other nodes, depending on the layer that node are located in. These inputs form an input vector  $\mathbf{X} = (x_1 \dots x_i \dots x_n)$ . The sequence of weights leading to the node form a weight vector  $\mathbf{W}_j = (w_{1j}, \dots, w_{ij}, \dots, w_{nj})$ , where  $w_{ij}$  represents the connection weight from  $i^{\text{th}}$  node in the preceding layer to this node.

The output of node  $j$ ,  $y_j$  is obtained by computing the value of function  $f$  with respect to the inner product of vector  $\mathbf{X}$  and  $\mathbf{W}_j$  minus  $b_j$ , where  $b_j$  is the threshold value, also called the bias, associated with this node. The following equation defines the operation :

$$y_j = f(\mathbf{X} \cdot \mathbf{W}_j - b_j) \quad (5.1)$$



**Fig. 5.2 A Schematic Diagram of a Typical  $j$ -th Node**

The function  $f$  is called an activation function. Its functional form determines the response of a node to the total input signal it receives. The most commonly used form of activation function is the *sigmoid function* given as:

$$f(x) = \frac{1}{1 + e^{-\beta x}} \quad (5.2)$$

This function is a continuous function that varies gradually between asymptotic values 0 and 1.

Where,  $\beta$  is the slope parameter, which adjusts the abruptness of the function as it changes between the two asymptotic values. Sigmoid functions are differentiable, which is an important feature of neural network theory. The sigmoid function is bounded, monotonic, non-decreasing function that provides a graded, nonlinear response. This function enables a network to map any nonlinear process. The popularity of the sigmoid function is partially attributed to the simplicity of its derivative that will be used during the training process.

### 5.1.2 Network Learning/Training

In order for an ANN to generate an output vector  $\mathbf{Y} = (y_1, \dots, y_i, \dots, y_n)$  that is as close as possible to the target vector  $\hat{\mathbf{Y}} = (\hat{y}_1, \dots, \hat{y}_i, \dots, \hat{y}_n)$ , a training process called learning, is employed to find optimal weight matrices  $\mathbf{W}$  and bias vectors  $\mathbf{B}$ , that minimize a predetermined error function that usually has the form :

$$E_f = \sum_P \sum_p (y_i - \hat{y}_i)^2 \quad (5.3)$$

Here  $\hat{y}_i$  is the component of desired output  $\hat{\mathbf{Y}}$ ,  $y_i$  is the corresponding ANN output,  $p$  is the number of output nodes, and  $P$  is the number of training patterns. The training is a process by which the connection weights of an ANN are adopted through a continuous process of stimulation by the environment in which the network is embedded. There are primarily two types of training – *supervised and unsupervised*. A supervised training algorithm requires an external teacher to guide the training processes. This typically implies that a large number of examples (or patterns) of inputs are required for training.

The inputs are cause variables of a system and outputs are the effect variables. This training procedure involves the iterative adjustment and optimization of connection weights and threshold values for each of the nodes. The primary goal of training is to minimize the error function by searching for a set of connection strengths and threshold values that cause the ANN to produce outputs that are equal or close to targets. After training has been accomplished, it is hoped that the ANN is then capable of generating reasonable results given



new inputs. In contrast, an unsupervised training algorithm does not involve a teacher. During training, only an input data set is provided to the ANN that automatically adopts its connection weights to cluster those input pattern into classes with similar properties. Most hydrological applications have utilized supervised training. In the present study supervised and Back-propagation algorithms have been adopted (ASCE,2000).

#### 5.1.4 Back-Propagation

Back propagation is perhaps the most popular algorithm for training ANNs. It is essentially a gradient descent technique that minimizes the network error function i.e., Eq. (5.3). Each input pattern of the training data set is passed through the network from the input layer to the output layer. The network output is compared to the desired target output, and an error is computed based on Eq. (5.3). This error is propagated backward through the network to each node and correspondingly the connection weights are adjusted based on equation:

$$\Delta w_{ij}(n) = -\epsilon \left( \frac{\partial E}{\partial w_{ij}} \right) + \alpha \Delta w_{ij}(n-1) \quad (5.4)$$

Where  $\Delta w_{ij}(n)$  and  $\Delta w_{ij}(n-1)$  are weight increments between node  $i$  and  $j$  during the  $n^{\text{th}}$  and  $(n-1)^{\text{th}}$  pass, or epoch. A similar equation is written for correction of bias values. In Eq. 5.4  $\epsilon$  and  $\alpha$  are called learning rate and momentum factor, respectively. The momentum factor can speed up training in very flat region of error surface and help prevent oscillations in the weights. A learning rate is used to increase the chance of avoiding the training process being trapped in local minima instead of global minima.

The back propagation algorithm involves two steps. The first is the forward pass, in which the effect of input is passed forward through the network to reach the output layer. After the error is computed, a second step starts backward through the network. The error at the output layers are propagated back towards the input layer with the weights modified according to the Eq. (5.4). Back propagation is the first order method based on the steepest gradient descent, with the direction vector being set equal to the negative of the gradient vector. Consequently the solution often follows a zigzag path while trying to reach a minimum error position, which may slow down the training process. It is also possible for the training process to be trapped in the local minimum despite the use of learning rate.

### **5.1.5 Advantage of ANNs**

ANNs offer certain advantages over the traditional rule-based system, i.e. conventional programming and knowledge-based expert systems. ANNs are preferable because of the following reasons:

1. They are weighted connection and massively parallel processing with fault tolerance in the sense that their performance degrades gracefully under adverse operating conditions and they can automatically learn from experience. This is called internal representation.
2. They have the generalization capability to learn complex patterns of inputs and provide meaningful solutions to problems even when input data contain errors, or are incomplete, or are not presented during training.
3. They are distribution free because no prior knowledge is needed about the statistical distribution of the classes in the data sources in order to apply the method for classification..
4. They take care of determining how much weight each data source should have in the classification, which remains a problem for statistical methods. The non-linear learning and smooth interpolation capabilities give the neural network an edge over standard computers and rule-based systems for solving certain problems.
5. They are able to recognize the relation between the input and output variables without knowing physical consideration.
6. They work well even when the training set contains noise and measurement errors. There is no need to make assumptions about the mathematical form of the relationship between input and output.

### **5.1.7 Limitation of ANN**

ANNs are unable to reason in a sequential or stepwise manner that results in precise conclusions. These restrictions could be critical when dealing with situations that demand exact answers and lucid justifications. Due to the difficulty in explaining, the only way to test the system for consistency and reliability is to monitor the output.

Back-propagation networks suffer from four main problems. The first problem is that network structuring is a versatile, intuitive, and highly solution-dependent trial-and-error task. The second is that the algorithm is slow in training, and convergence is very sensitive to the initial set of weights. The third is that training can be trapped in local minima. The fourth is that the design of an optimum network configuration for a given problem is a non-guided or trial-and-error process that does not guarantee adequate generalization.

## **5.2 Development of ANN Models**

### **5.2.1 Data Partition**

Whole data of rainfall, runoff, temperature and sediment yield have been divided into three sets: training, testing and validation. The training set was used to adjust the connection weights, whereas the testing set was used to check the performance of the model at various stages of training and to determine when to stop training to avoid over-fitting. The validation set was used to estimate the performance of the trained network in the deployed environment.

#### ***(a) Rainfall-Runoff Model***

Daily rainfall, runoff and temperature data were used to develop the various runoff prediction models. The data from 1995 to 1999 are available at the different stations in the watershed and they were divided into three sets: training (70%), testing (20%) and validation (10%).

#### ***(b) Sediment Yield Prediction Model***

The observed data of daily sediment yield and runoff from the year 2001 to 2003 were used for the model development. As no continuous data series were available at the site, only wet season data were adopted. In the year 2001, data from July to December are available, similarly data of July to October and July to September were found properly recorded in the year 2002 and 2003, respectively. So all together 387 data set were used after partitioning them into training (70%), testing (20%) and validation (10%) sets.

## 5.2.2 Selection of Input and Output Variable and Their Normalization

### (a) Selection of Input and Output Variable

The selection of an appropriate input vector,  $\mathbf{X}$  that will allow an ANN to successfully map to the desired output vector,  $\mathbf{Y}$  is not trivial task. The selection of training data that represents the characteristics of a watershed and meteorological pattern is extremely important in modelling. The training data should be large enough to contain the characteristics of the watershed and to accommodate the requirement of ANN architecture. If the information included in the training set is insufficient, an increase in the complexity of network will not enable the network to generate the pattern in the physical phenomena. On the contrary, an increase in the complexity of the models might mislead the modeller to over fit the training data and lead to the poor forecasts.

In the present study, data of daily rainfall, runoff and temperature at various locations as shown in Fig.3.2 have been selected and presented in the Table 5.1. The details of rainfall, runoff and temperature data are reported in Appendix A, B and C.

**Table 5.1 Selected Input and Output Variables**

#### (a) Rainfall-Runoff Modelling

S.N.	Variables	Duration	Location
1	Daily Rainfall "R <sub>1</sub> "	1995 - 1999	Ilam tea estate
2	Daily Rainfall "R <sub>2</sub> "	1995 - 1999	Sokatim
3	Daily Rainfall "R <sub>3</sub> "	1995 - 1999	Kanyam
4	Daily Temperature "T <sub>1</sub> "	1995 - 1999	Ilam tea estate
5	Daily Temperature "T <sub>2</sub> "	1995 - 1999	Sokatim
6	Daily Temperature "T <sub>3</sub> "	1995 - 1999	Kanyam
7	Daily Runoff "Q"	1995 - 1999	Mainachuli

#### (b) Sediment Yield Modelling

S.N.	Variables	Duration	Location
1	Daily Runoff "Q"	June 29 -Dec 30 (2001)	Mainachuli
2	Daily Suspended SedimentYield "Q <sub>s</sub> "	July 2 -Sep 25 (2002)	
		June 23 -Oct 17 (2003)	

Before deciding the input parameters, autocorrelation and cross correlation analysis have been carried out considering different lag times. Analysis indicates as shown in Fig. 5.3 that for this particular watershed the rainfall data and previous time steps runoff having longer lag time have poor correlation with runoff. So rainfall data having time steps of t, t-1, and t-2 ; and runoff data of time step t-1 and t-2 are considered for developing various runoff prediction models and performance of these models were tested using different performance evaluation criteria and effect of various inputs were investigated. Similarly runoff has very poor correlation with temperature, (Fig. 5.4) but sometimes though there is no high correlations between individual data type, when combined with different sets may give good result. So in some of the models, temperature has also been considered as one of the inputs. As per the result of correlation analysis, temperatures recorded at Ilam tea estate, Saktim and Kanyam have been taken with their time step t-26, t-8 and t-9, respectively.

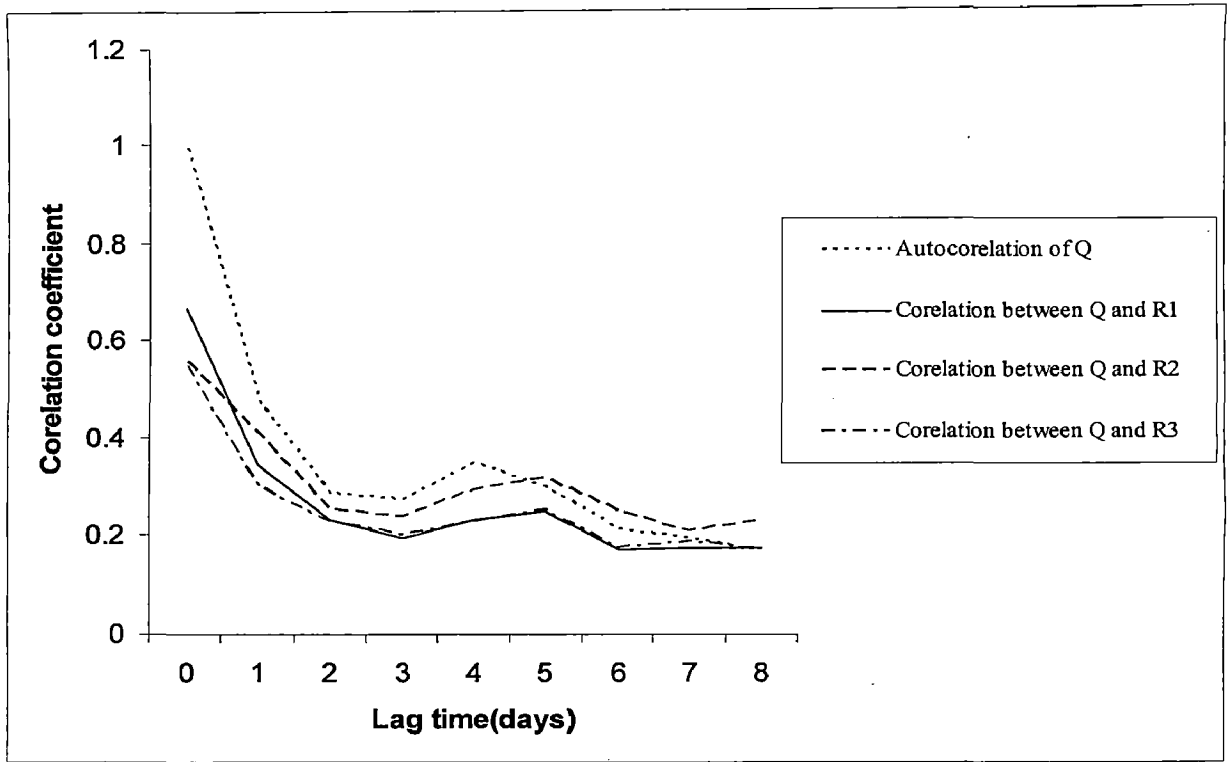
The autocorrelation and cross correlation analysis have also been carried out between sediment yield of time step 't' and sediment yield and runoff of previous time steps (t-1, t-2, t-3 etc.(Fig. 5.5). As in the case of runoff, runoff and sediment yield of longer lag time have poor relation with sediment yield. So only time steps of t-1 has been adapted. Based on the above analysis different combinations of inputs that have been considered for developing the models are given in Table 5.2 and 5.3.

#### ***(b) Normalization***

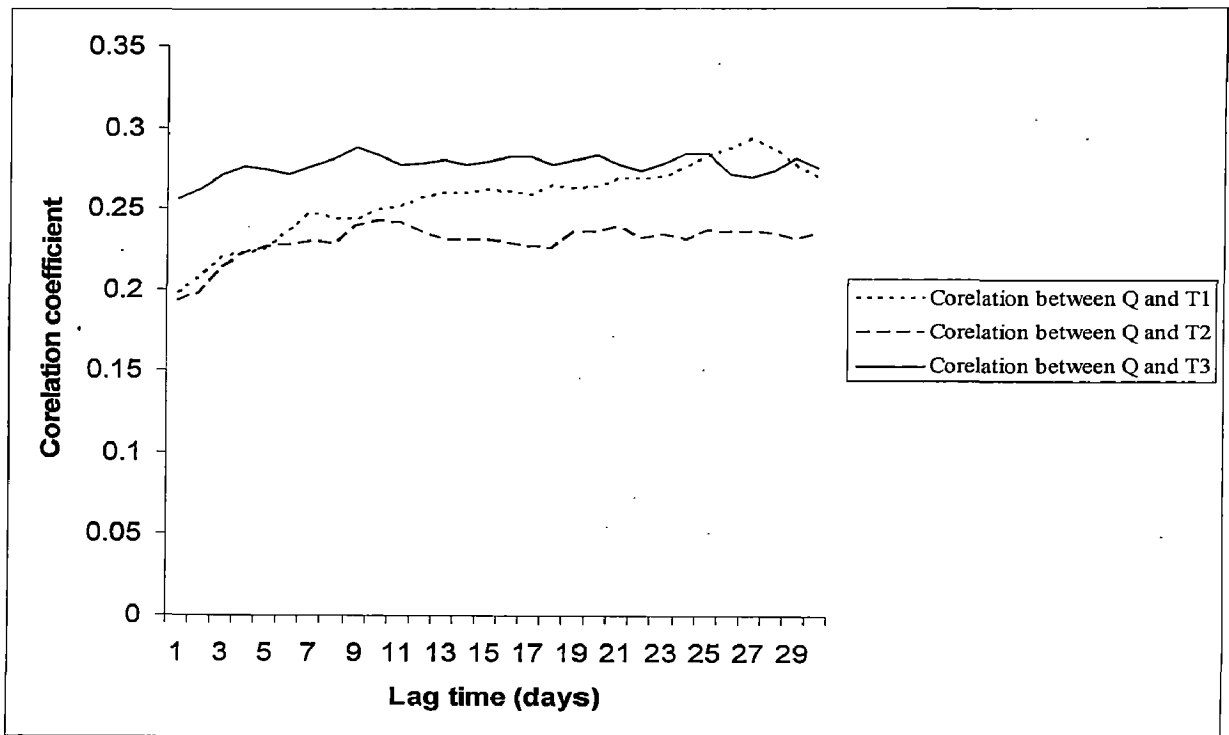
Normalization of input and output data is the important step in ANN model development. It is mentioned that the sigmoid function can take the values ranging in the (0, 1) domain. There are different normalization procedures; the followings method converts the original data series into the range (0.0 and 1.0).

$$X_{Ni} = (X_i - X_{\min}) / (X_{\max} - X_{\min}) \quad (5.5)$$

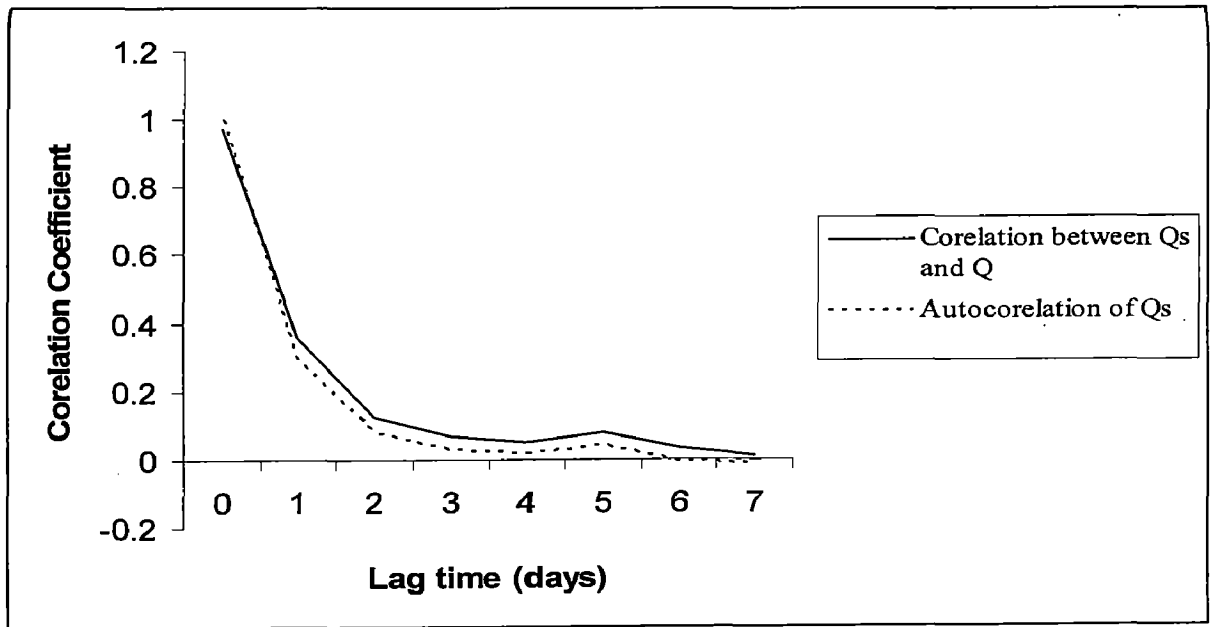
Where:  $X_N$  is the normalized value,  $X$  is the original variable;  $X_{\min}$  is the minimum value of variable,  $X_{\max}$  is the maximum value of variable ; and  $i$  is an index representing the number of data points.



**Fig. 5.3 Autocorrelation of Runoff and Cross Correlation between Runoff and Rainfall.**



**Fig. 5.4 Correlation between Runoff and Temperature**



**Fig. 5.5 Autocorrelation of Sediment Yield and Cross Correlation between Sediment Yield and Runoff.**

### 5.2.3 Designing an ANN

There are no fixed rules for developing an ANN, even though a general framework can be followed based on previous successful application in engineering. In the present study a model based on a feed forward neural network; with sigmoid function as the transfer function, and single hidden layer has been used. ANN model implementation was carried out using the Neural Power software. The connection weights, and number of neurons in the hidden layer, which can be interpreted as the model parameter, were adjusted during the training process through the minimization of mean square error (MSE) using error back propagation algorithm. For each ANN configuration the training procedure was repeated starting with independent initial condition ultimately ensuring best performing network. The trend of decrease in the MSE in training and testing sets was used to decide the optimal learning. The training was stopped when the MSE over the testing set started rising instead of reducing even though the MSE over the training sets continued to decrease. This is the indication of the network getting overstrained; as such an ANN model would perform very well in the training period but would fail to maintain that level of performance when applied to different data sets.

The numbers of nodes in the input layers are equal to the number of input variables, nodes in the output layer is one as the model provides single output i.e. runoff. The number of nodes in the hidden layer, which are responsible for capturing the (or mapping) the complex, dynamic, non-linear rainfall- runoff process is a trial and error job .To begin with the number of nodes in the hidden layer has been taken as provided in default setting of Neural power software. The sensitivity of error of each weight was checked and nodes in the hidden layer were increased or decreased if necessary. Many single hidden -layer ANN architecture were investigated before deciding the final one. The value of learning and momentum correction factors of 0.1 and 0.4 respectively employed during the training.

#### **5.2.4 Performance and Evaluation Criteria of Models**

The performance of a model can be evaluated in terms of accuracy, consistency and versatility. The term accuracy refers to the ability of the model to reduce the calibration error consistency is used for representing the characteristics of the model whereby the level of accuracy and estimate of the parameters values persists through different samples of the data. A versatile model is defined as the model which is accurate and consistent when used for different application.

A variety of verification criteria, which could be used for the evaluation, and inter-comparison of different models may be grouped as graphical and numerical performance indicator. Of the various graphical indicators following indicators are used in the present study (Sajikumar et. al. 1999).

- (i) A linear scale plot of the simulated and observed daily runoff.
- (ii) Double Mass curve plots of simulated and observed flows.
- (iii) A scatter plot of simulated versus observed flows.

Numerical indicators are the root mean square error (RMSE),  $R^2$  efficiency (Nash and Sutcliffe, 1970) and Coefficient of correlation (CC).



(a) **Root Mean Square Error:** It yields a residual error in terms of mean square error, expressed as:

$$\text{RMSE} = \sqrt{\text{residual variance} / n} = \left[ \sum_{j=1}^n (Y_j - \hat{Y}_j)^2 / n \right]^{1/2} \quad (5.6)$$

$Y$  and  $\hat{Y}$  are the estimated and observed values and  $n$  is the number of observations.

(b) **Correlation Coefficient (CC):** It is expressed as:

$$\text{CC} = \sum_{j=1}^n \left\{ (\hat{Y}_j - \bar{\hat{Y}}) (\bar{Y} - Y_j) \right\} \left\{ \left( \sum_{j=1}^n (\hat{Y}_j - \bar{\hat{Y}})^2 \right) \left( \sum_{j=1}^n (Y_j - \bar{Y})^2 \right) \right\}^{1/2} \quad (5.7)$$

Where  $\bar{Y}$  and  $\bar{\hat{Y}}$  are mean of estimated and observed values.

(c) **Coefficient of Efficiency (CE)**

Based on the standardization of residual variance with initial variance it is expressed as :

$$R^2 = \{ 1 - \text{residual variance} / \text{initial variance} \} \times 100 \quad (5.8)$$

$$R^2 = \left\{ 1 - \frac{\sum_{j=1}^n (\hat{Y}_j - Y_j)^2}{\sum_{j=1}^n (\hat{Y}_j - \bar{\hat{Y}})^2} \right\} \times 100$$

### 5.2.5 Rainfall-Runoff Model

The runoff prediction models have been developed for the runoff measurement site "Mainachuli". All together eleven models have been developed considering different combination of rainfall, temperatures and previous time step runoff. Performance of various models is presented in the Table 5.2 and the scatter plots indicating error line are given in Fig. 5.9 to Fig. 5.16.

Table 5.2 depicts that performance of Model 4 is better than the other models. Therefore details of model 4 are discussed below:



**Model 4:**

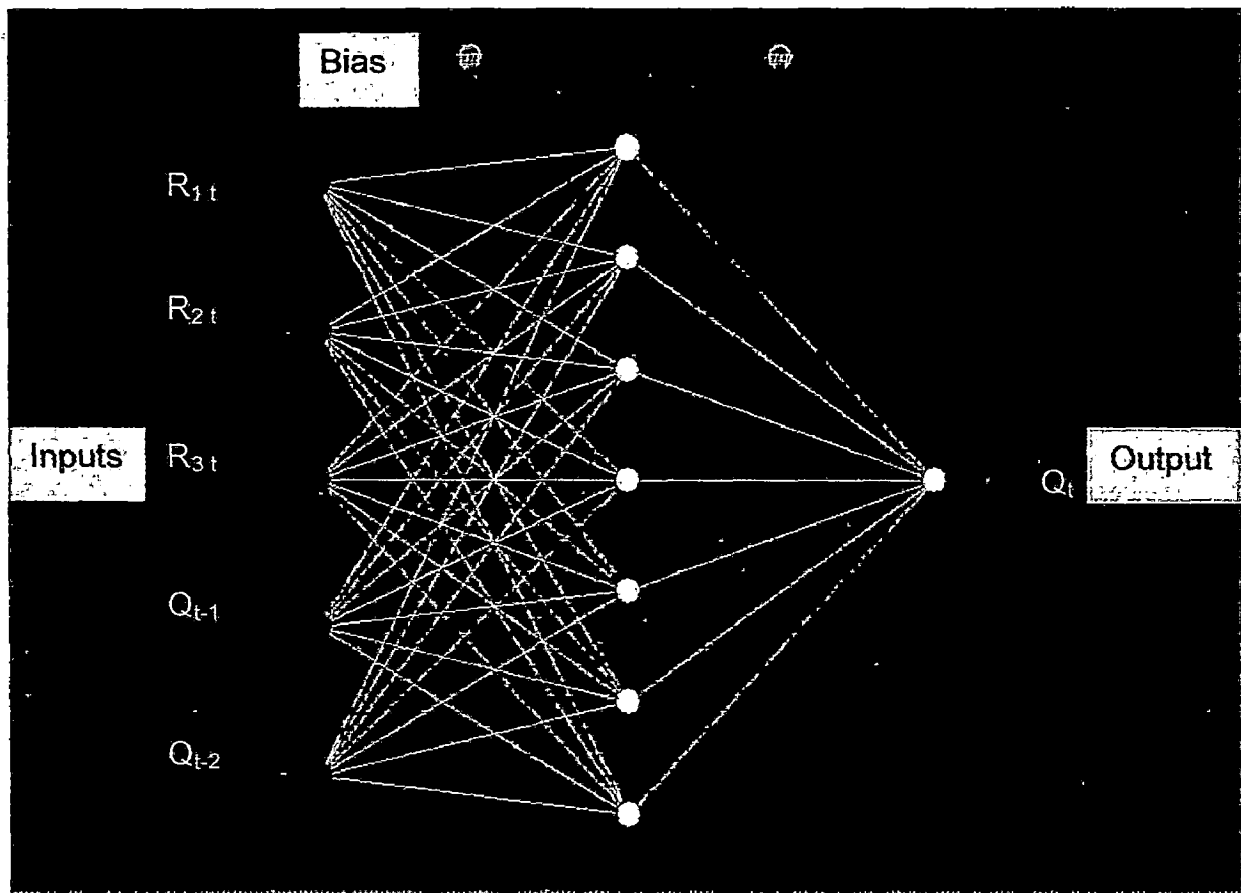
$$Q_t = f(Q_{t-1}, Q_{t-2}, R_{1t}, R_{2t}, R_{3t})$$

The model 4 was constituted to simulate the current runoff from the known rainfall of three rain gauge stations of time step “t” (i.e.  $R_{1t}, R_{2t}, R_{3t}$ ) and previous days runoff of time step t-1 ( $Q_{t-1}$ ) and time step t-2 ( $Q_{t-2}$ ).

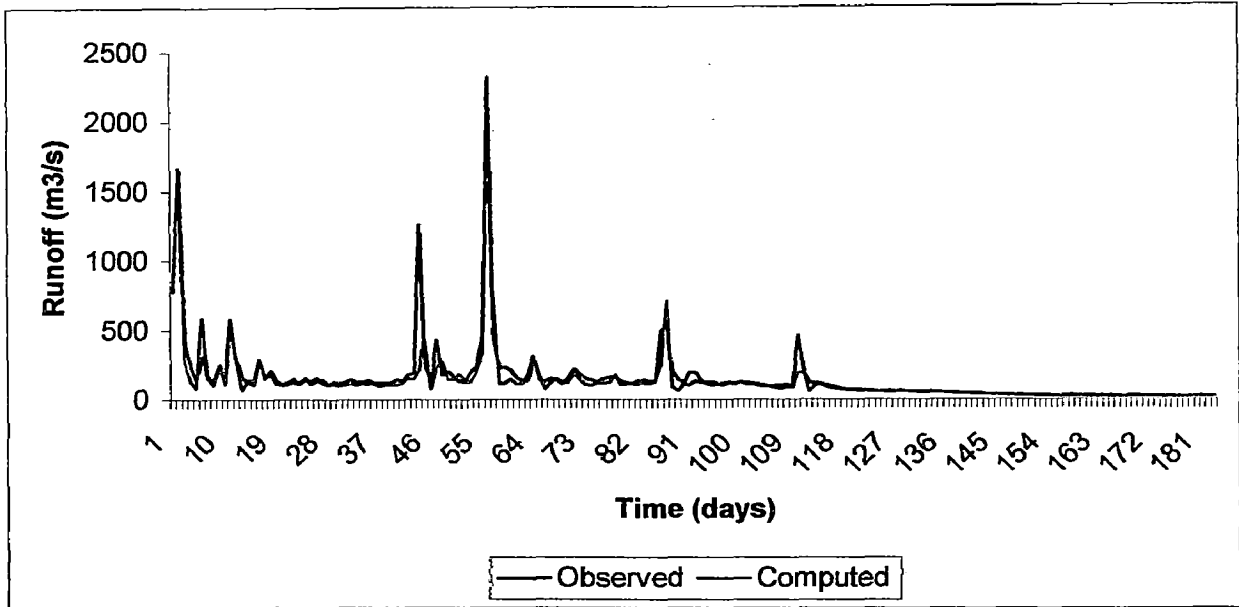
The model architecture was designated as 5-7-1, indicating 5 input neurons, one hidden layer with 7 neurons and 1-output neurons.

In this model, an RMSE error of 103.67 m<sup>3</sup>/sec and Nash co-efficient of 0.82 were obtained. Average correlation co-efficient ‘CC’ was 0.91. Figure 5.7 shows the plot of validation curve between observed data and ANN predicted data. The scatter plot of predicted verses observed runoff are also shown in Fig. 5.9 and corresponding mass curve plots of the observed and predicted runoff are shown in Fig.5.8.

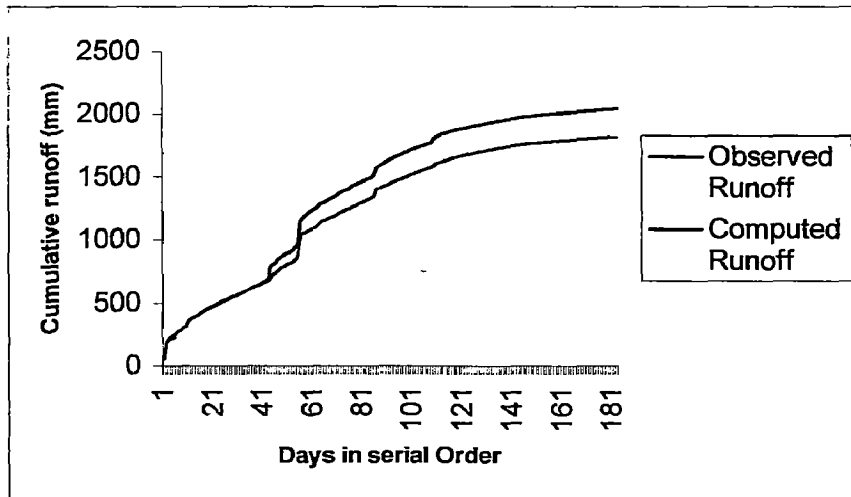
Details of weighting factors and bias are given in Table 5.4.



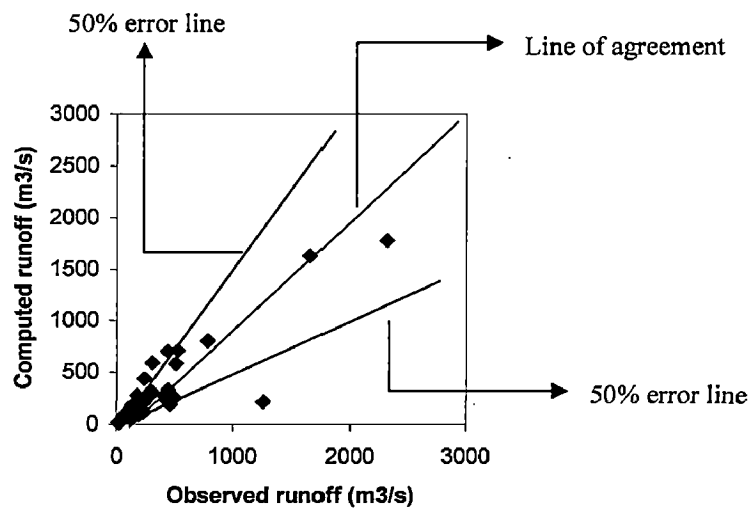
**Fig. 5.6 ANN Architecture of Model 4**



**Fig. 5.7 Validation of Computed Runoff with Observed one**



**Fig. 5.8 Mass Curve of the Computed Runoff and Observed Runoff for Validation Data**



**Fig. 5.9 Scatter Plot of Computed Runoff against Observed Runoff for Validation Data**

**Table 5.2 Runoff prediction models**

Model No.	Model Descriptions	Number of Nodes						Learning /Training Data						Validation Data						
		I	H	O	R <sup>2</sup>	CC	iteration	€	α	Transfer function	RMSE (m <sup>3</sup> /s)	R <sup>2</sup>	CC	R <sub>R</sub>	Z.	RMSE (m <sup>3</sup> /s)	R <sup>2</sup>	CC	R <sub>R</sub>	Z.
1	$Q_t = f(R_{1t}, R_{2t}, R_{3t})$	3	3	1	.68	.83	4000	.4	.4	sigmoid	99.10	.68	.83	.50	48	151.05	.60	.89	.50	48
2	$Q_t = f(Q_{t-1}, R_{1t}, R_{2t}, R_{3t})$	4	4	1	.85	.92	4000	.4	.4	sigmoid	68.77	.85	.92	.77	96	115.75	.76	.90	.77	96
3	$Q_t = f(Q_{t-1}, R_{1t}, R_{2t}, R_{3t}, T_{1t-26}, T_{2t-9}, T_{3t-8})$	7	6	1	.88	.94	4500	.4	.4	sigmoid	61.54	.88	.94	.85	94	130.76	.81	.91	.85	94
4	$Q_t = f(Q_{t-1}, Q_{t-2}, R_{1t}, R_{2t}, R_{3t})$	5	7	1	.88	.94	4500	.4	.4	sigmoid	61.92	.88	.94	.89	94	103.67	.82	.91	.89	94
5	$Q_t = f(Q_{t-1}, R_{1t}, R_{2t}, R_{3t}, R_{1t-1}, R_{2t-1}, R_{3t-1})$	7	7	1	.92	.96	5000	.4	.4	sigmoid	49.56	.92	.96	.96	93	139.22	.66	.83	.96	93
6	$Q_t = f(Q_{t-1}, Q_{t-2}, R_{1t}, R_{2t}, R_{3t}, R_{1t-1}, R_{2t-1}, R_{3t-1})$	8	6	1	.86	.93	5000	.4	.4	sigmoid	65.58	.86	.93	.86	92	130.48	.71	.85	.86	92
7	$Q_t = f(Q_{t-1})$	1	3	1	.36	.60	3200	.4	.4	sigmoid	142.19	.36	.60	.73	71	192.29	.35	.62	.73	71
8	$Q_t = f(R_{1t}, R_{2t}, R_{3t}, T_{1t-26}, T_{2t-9}, T_{3t-8})$	6	5	1	.55	.81	4500	.4	.4	sigmoid	160.06	.55	.81	.59	56	129.35	.71	.89	.59	56

**Table 5.3 Sediment yield prediction models**

Model No.	Model Descriptions	Number of Nodes						Learning /Training Data						Validation Data						
		I	H	O	R <sup>2</sup>	CC	iteration	€	α	Transfer function	RMSE (t/day)	R <sup>2</sup>	CC	R <sub>R</sub>	Z.	RMSE (t/day)	R <sup>2</sup>	CC	R <sub>R</sub>	Z.
1	$Q_{st} = f(Q_t)$	1	2	1	.90	.95	400	.4	.4	sigmoid	18526	.90	.95	1.18	42	19058	.93	.97	1.18	42
2	$Q_{st} = f(Q_t, Q_{t-1})$	2	2	1	.85	.93	500	.4	.4	sigmoid	16012	.85	.93	1.09	40	26052	.90	.96	1.09	40
3	$Q_{st} = f(Q_t, Q_{t-1})$	3	2	1	.78	.87	500	.4	.4	sigmoid	14127	.78	.87	1.17	37	34345	.80	.90	1.17	37
4	$Q_{st} = f(Q_t, Q_{t-1}, Q_{st-1})$	2	2	1	.77	.89	600	.4	.4	sigmoid	15234	.77	.89	.93	42	26925	.87	.94	.93	42

Where R<sub>R</sub> is recovery ratio i.e., ratio of cumulative predicted runoff/ sediment yield and cumulative observed runoff/sediment yield  
 Z is the percentage data falling under 50% error line

**Table 5.4 Details of Weighting Factors and Bias of Runoff Prediction Model**

Links	Weights	Links	Weights/bias
N1L1 - N1L2	-12.64	N4L1 - N5L2	-9.94
N1L1 - N2L2	1.29	N4L1 - N6L2	22.48
N1L1 - N3L2	0.66	N4L1 - N7L2	-4.53
N1L1 - N4L2	37.36	N5L1 - N1L2	1.73
N1L1 - N5L2	-12.51	N5L1 - N2L2	0.85
N1L1 - N6L2	0.76	N5L1 - N3L2	-1.57
N1L1 - N7L2	9.67	N5L1 - N4L2	0.12
N2L1 - N1L2	-0.88	N5L1 - N5L2	22.59
N2L1 - N2L2	-2.22	N5L1 - N6L2	6.02
N2L1 - N3L2	-56.80	N5L1 - N7L2	1.47
N2L1 - N4L2	98.10	B1 - N1L2	0.26
N2L1 - N5L2	8.85	B1 - N2L2	2.85
N2L1 - N6L2	17.74	B1 - N3L2	1.07
N2L1 - N7L2	-3.37	B1 - N4L2	0.91
N3L1 - N1L2	-1.87	B1 - N5L2	-0.20
N3L1 - N2L2	2.16	B1 - N6L2	-0.83
N3L1 - N3L2	5.23	B1 - N7L2	0.80
N3L1 - N4L2	8.40	N1L2 - N1L3	-1.44
N3L1 - N5L2	-3.58	N2L2 - N1L3	6.66
N3L1 - N6L2	1.78	N3L2 - N1L3	-1.59
N3L1 - N7L2	-4.58	N4L2 - N1L3	-0.31
N4L1 - N1L2	0.53	N5L2 - N1L3	0.20
N4L1 - N2L2	2.40	N6L2 - N1L3	-0.31
N4L1 - N3L2	9.32	N7L2 - N1L3	-0.61
N4L1 - N4L2	-10.30	B2 - N1L3	0.95

Where L1, L2 and L3 represent the input, hidden and output layer. N1, N2,.....N7 represent the nodes in each layer as the case may be.

### 5.2.6 Sediment Yield Prediction Model:

The sediment yield prediction models have been developed for the sediment yield at site“Mainachuli”. All together four ANN models have been developed considering different combination of observed runoff at time t and t-1( $Q_t$ ,  $Q_{t-1}$ ) and observed sediment yield at time t-1 ( $Q_{st-1}$ ). Other procedure is same as in runoff prediction model. Out of the above four models, *model 1* provides better result and the same is discussed below in detail.

#### Model 1:

$$Q_{st} = f(Q_t)$$

This model has been developed to simulate the suspended sediment concentration at Mainachuli station in the Kankaimai River considering present time runoff ( $Q_t$ ) as inputs. Figure 5.10 depicts the architecture of the sediment yield ANN model 1 showing the input and output parameters. The model architecture was designated as (1 - 2 - 1), indicating one input neuron, two hidden layers and one-output neurons.

In this model, an RMSE error of 19058 t/day and Nash co-efficient of 0.933 were obtained. Average correlation co-efficient ‘CC’ was 0.968. Figure 5.11 shows the plot of validation curve between observed data and ANN predicted data. The scatter plot of predicted verses observed sediment yield are also shown in Fig. 5.13 and corresponding mass curve plots of the observed and predicted sediment yield are shown in Fig.5.12. Details of weighing factors and bias are given in Table 5.5.

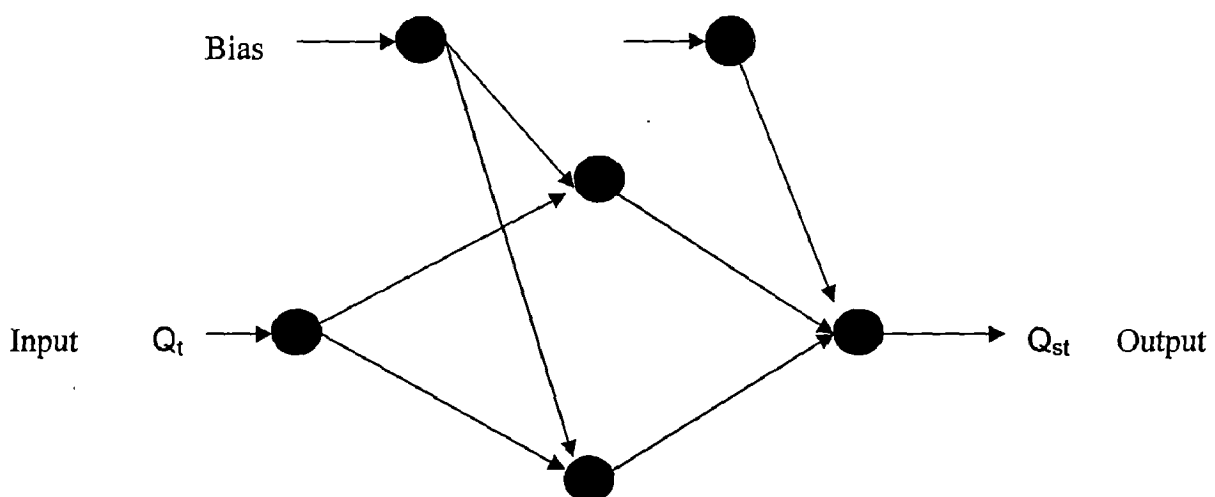
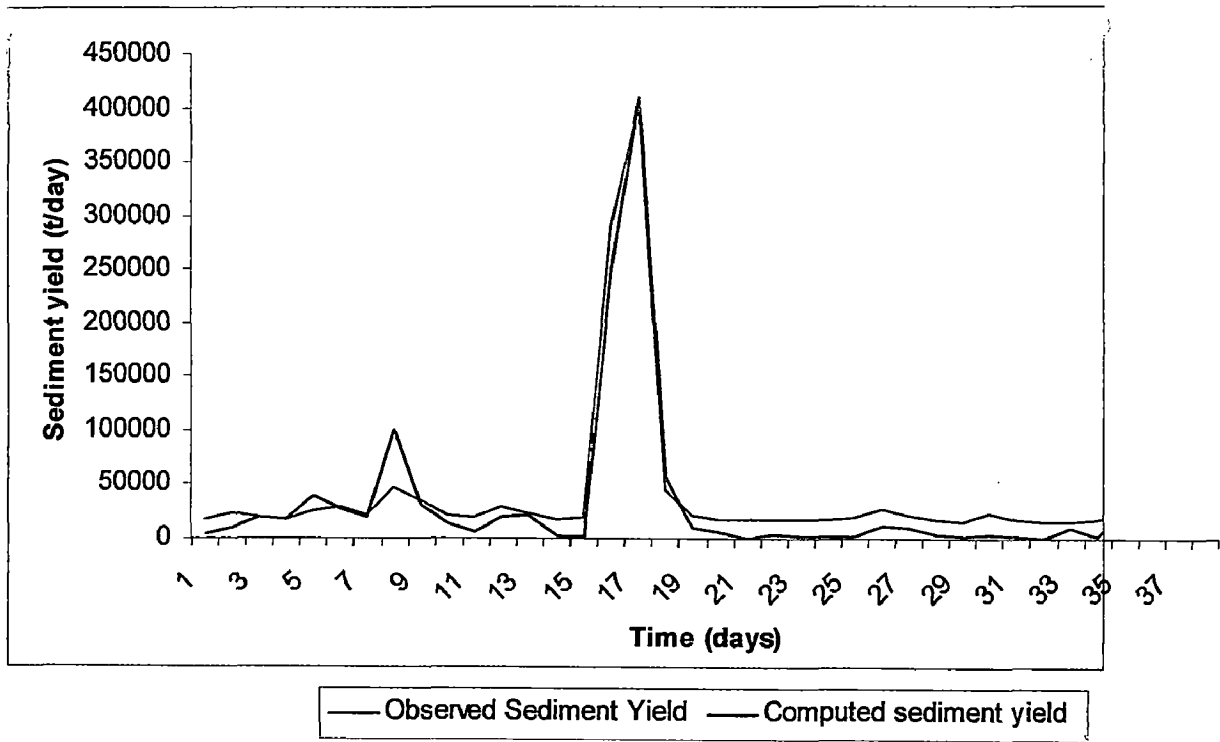
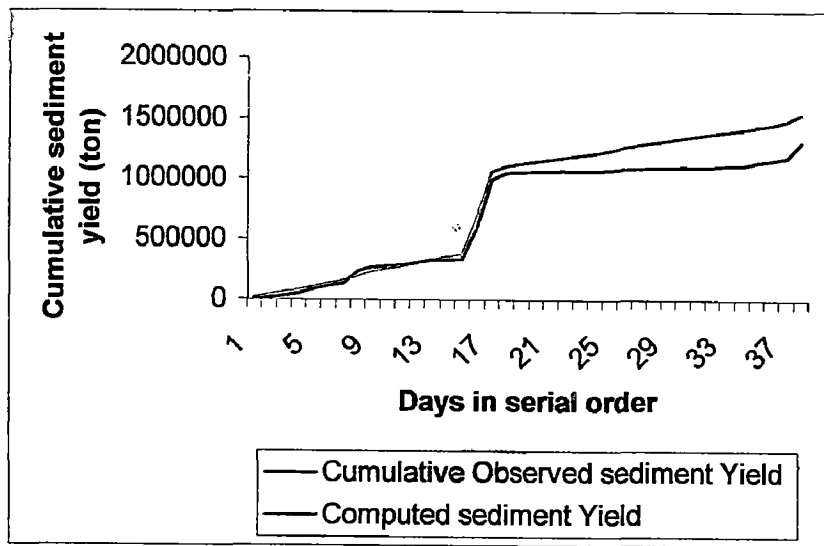


Fig. 5.10 ANN Architecture of Sediment Yield Prediction Model 1

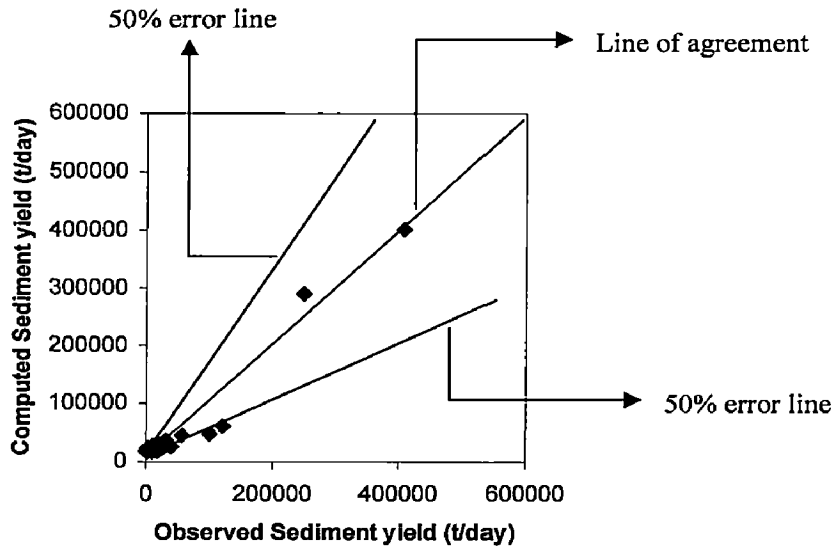


**Fig. 5.11 Validation of Computed Sediment Yield with Observed one**



**Fig. 5.12 Mass Curve of the Computed Sediment Yield and Observed Sediment Yield for Validation Data**





**Fig.5.13 Scatter plot of Computed Sediment Yield against Observed Sediment Yield for Validation Data**

**Table 5.5 Details of Weighting Factors and Bias of Sediment Yield Prediction Model**

Links	Weights
N1L1 - N1L2	1.14
N1L1 - N2L2	-2.36
B1 - N1L2	0.58
B1 - N2L2	-0.37
N1L2 - N1L3	1.23
N2L2 - N1L3	-2.43
B2 - N1L3	1.08

## CHAPTER-6

# DEVELOPMENT OF REGRESSION MODEL

---

### 6.1 Introduction

The regression analysis may be broadly defined as the analysis of relationship between several independent or predictor variables as a dependent or criterion variable. It is one of the most widely used statistical tools because it provides a simple method for establishing a functional relationship among variables. The general computational problem that needs to be solved in multiple regression analysis is to fit a straight line to a number of points. In the simplest case- one dependent and one independent variable one can visualize this in scatter plot. The goal of linear regression procedures is to fit a straight line through the points. Specifically, a line is computed so that the squared deviations of the observed points from that line are minimized. Thus, this general procedure is sometimes also referred to as least squares estimation.

### 6.2 The Regression Equation

In linear regression, the independent variables are combined to form a linear equation of the form  $\beta_0 x_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$ . Any observation of the random variable Y, say  $y_i$ , deviates from the systematic component of the model by an amount denoted by  $\varepsilon_i$ , so that the model equation can be written.

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon_i \quad (6.1)$$

where  $y_i$  is the dependent variable and  $x_1, x_2, \dots, x_n$  are independent variable. The parameters  $\beta_0, \beta_1, \beta_2, \dots, \beta_n$  are regression coefficients and are determined from the data.

The standardized residuals  $\varepsilon_i$  have zero mean and unit standard deviation.

Different performance evaluation criteria as discussed in section 5.2.4 are used to compare the performance of the different models.

### 6.3 Rainfall-Runoff Model

For the regression analysis of Runoff prediction different combinations of input data as mentioned in the Table 5.2 are considered and equation for runoff model is written. For example for Model No 4, the equation may be written as;

$$Q_t = \beta_0 + \beta_1 R_{1t} + \beta_2 R_{2t} + \beta_3 R_{3t} + \beta_4 Q_{t-1} + \beta_5 Q_{t-2} \quad (6.2)$$

The regression analysis involves determining the parameters  $\beta_0, \beta_1, \beta_2, \beta_3, \text{ and } \beta_4$  with the help of available observed data set. Ninety percent data set have been used for calibration and ten percent data set are reserved for validation purpose.

Regression analysis has been carried in MS Excel 2003 and the result are reported in Table 6.1.

$$Q_t = 0.495 + 0.288R_{1t} + 0.018R_{2t} + 4.509R_{3t} + 2.049Q_{t-1} + 0.873Q_{t-2} \quad (6.3)$$

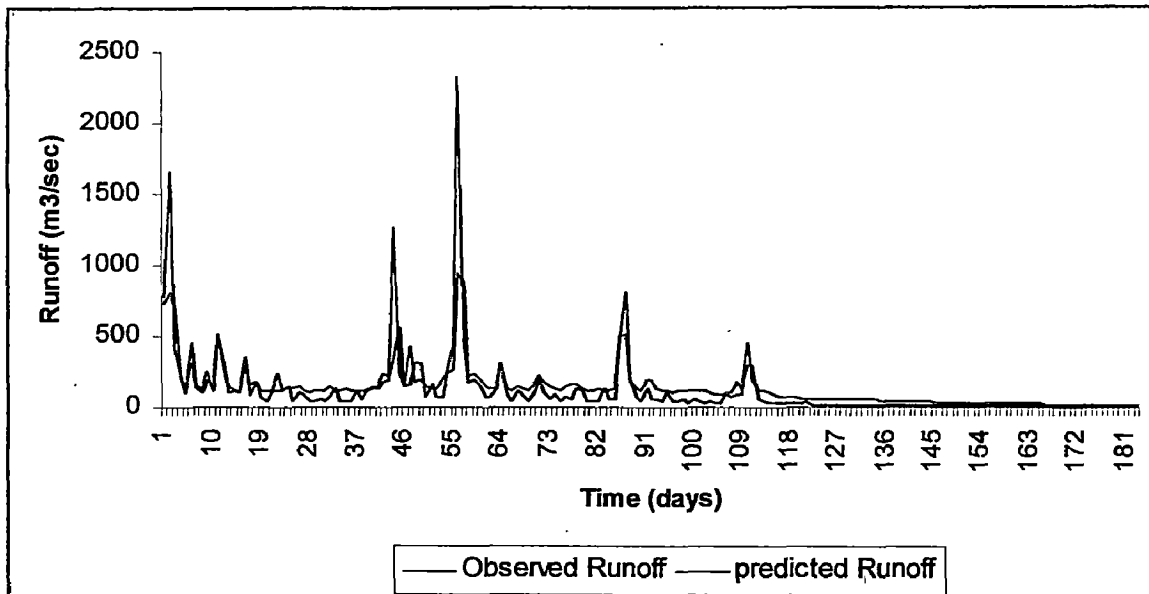
Various models have been validated with the unused sets of data and the predicted runoff value were plotted together with observed runoff data as shown in the Fig. 6.1 for Model 4. Similar cumulative plot of observed and predicted runoff; and scatter plot of predicted runoff against observed were plotted for Model 4 and presented in Fig. 6.2 and Fig. 6.3, respectively.

**Table 6.1 Performance of various regression models for runoff prediction**

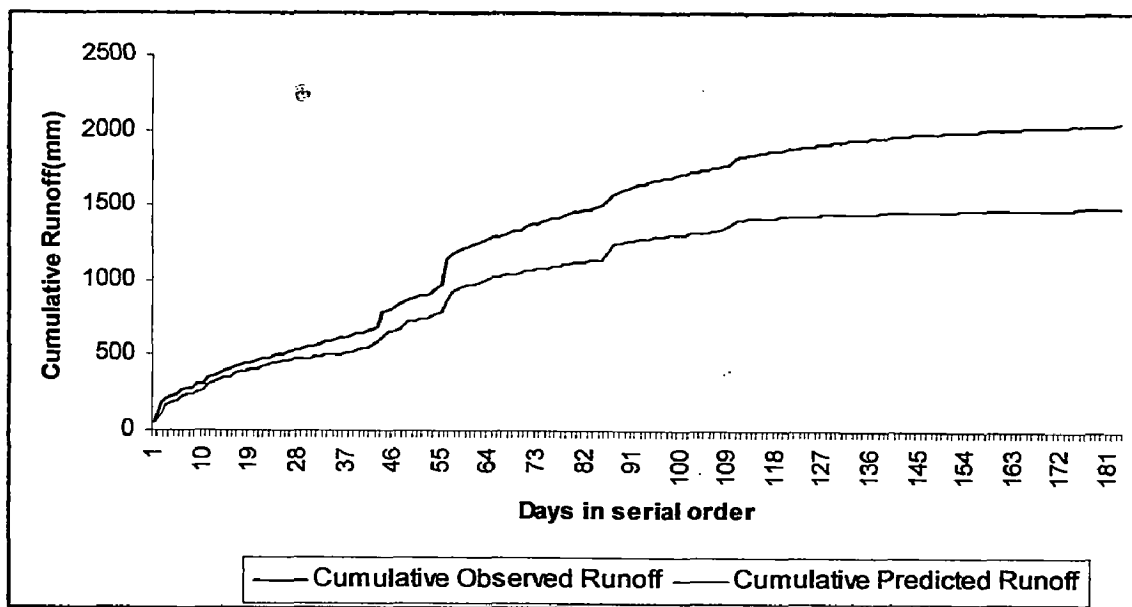
Model No	Regression Eq <sup>n</sup> .	Calibration			Validation				
		RMSE (m <sup>3</sup> /s)	R <sup>2</sup>	CC	RMSE (m <sup>3</sup> /s)	R <sup>2</sup>	CC	R <sub>R</sub>	Z.
1	$Q_t = 34.32 + 5.14R_{1t} + 2.63R_{2t} + 0.85R_{3t}$	126.7	0.493	0.70	162.95	0.54	0.81	0.57	41
2	$Q_t = 1.05 + 0.29Q_{t-1} + 4.5R_{1t} + 2.06R_{2t} + 0.88R_{3t}$	116.43	0.57	0.756	158.10	0.57	0.78	0.72	36
3	$Q_t = 33 + 0.295Q_{t-1} + 4.53R_{1t} + 2.08R_{2t} + 0.88R_{3t} + 1.54T_{1t-26} - 3.46T_{2t-9} + 1.28T_{3t-8}$	117.20	0.57	0.76	157.48	0.57	0.78	0.73	55
4	$Q_t = 0.49 + 0.29Q_{t-1} + 0.02Q_{t-2} + 4.5R_{1t} + 2.05R_{2t} + 0.87R_{3t}$	116.47	0.57	0.76	157.45	0.57	0.78	0.62	37
5	$Q_t = 2.97 + 0.33Q_{t-1} + 4.53R_{1t} + 2.04R_{2t} + 0.98R_{3t} + 0.21R_{1t-1} + 0.64R_{2t-1} - 1.11R_{3t-1}$	114.89	0.58	0.77	158.84	0.56	0.77	0.75	40
6	$Q_t = 2.71 + .32Q_{t-1} + .008Q_{t-2} + 4.53R_{1t} + 2.03R_{2t} + .97R_{3t} + .22R_{1t-1} + .64R_{2t-1} - 1.11R_{3t-1}$	114.95	0.58	0.76	158.55	0.56	0.77	0.75	40
7	$Q_t = 34.32 + .487 Q_{t-1}$	155.50	0.23	0.48	226.33	0.10	0.384	0.72	70
8	$Q_t = 18.61 + 5.12R_{1t} + 2.53R_{2t} + .78R_{3t} + 3.42T_{1t-26} + 3.78T_{2t-9} + 3.78T_{3t-8}$	127.05	0.49	0.70	159.89	0.60	0.81	0.61	47

**Table 6.2 Performance of various regression models for sediment yield prediction**

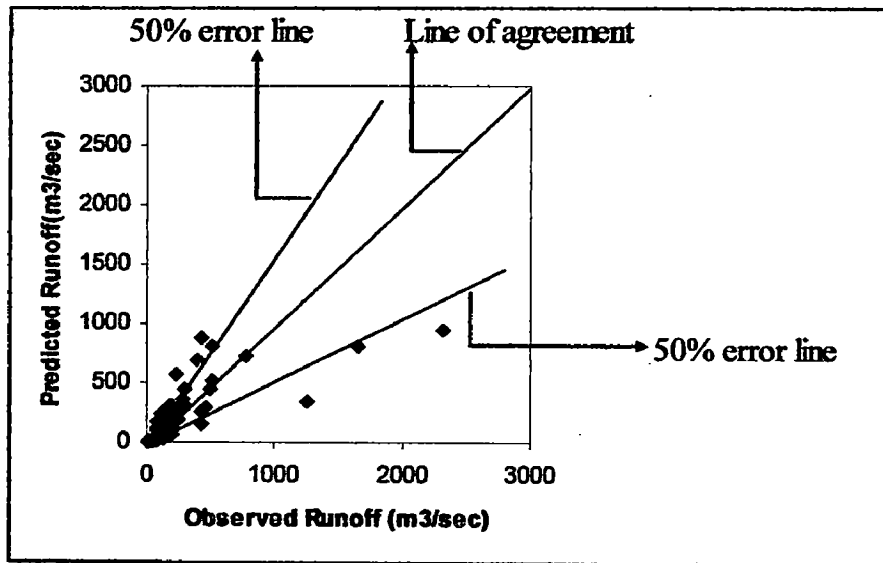
Model No	Regression Eq <sup>n</sup> .	Calibration			Validation				
		RMSE (t/day)	R <sup>2</sup>	CC	RMSE (t/day)	R <sup>2</sup>	CC	R <sub>R</sub>	Z.
1	$Q_{st} = -21594.6 + Q_t$	36832	0.82	0.91	106607	-0.85	0.97	1.65	29
2	$Q_{st} = -20548.35 + 220.51Q_t + 0.138Q_{st-1}$	35564	0.83	0.91	95101	-0.50	0.97	1.61	19
3	$Q_{st} = -22408.91 + 234.73Q_t + 11.38Q_{t-1}$	36797	0.82	0.91	104990	-0.83	0.97	1.66	19
4	$Q_{st} = -14005.97 + 227.30Q_t - 71.43Q_{t-1} + 0.366Q_{st-1}$	34125	0.85	0.92	83770	-0.16	0.97	1.46	10



**Fig. 6.1 Validation of Computed Runoff with Observed one**



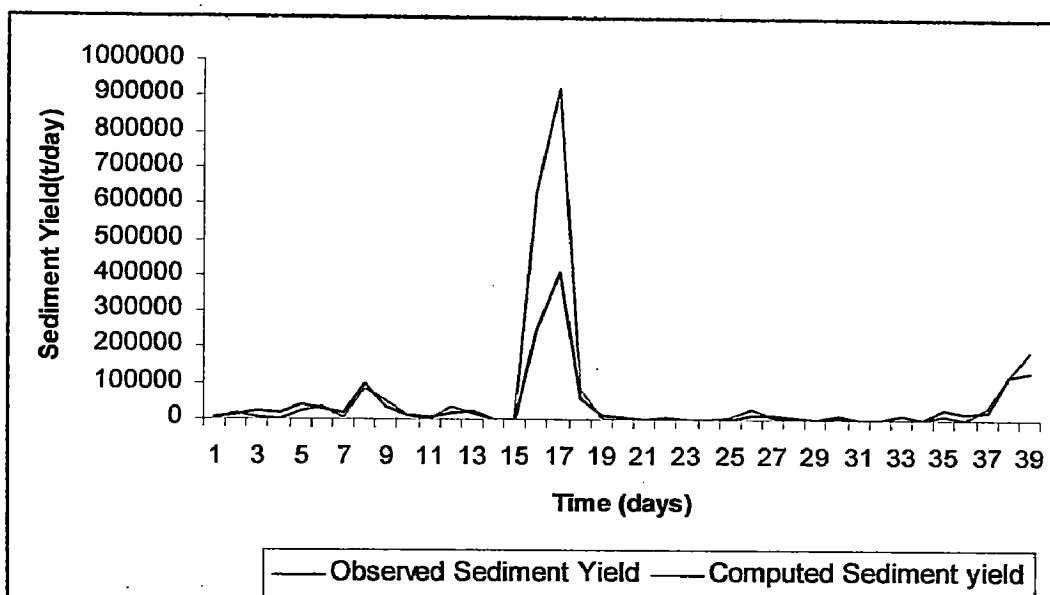
**Fig. 6.2 Mass Curve of Computed Runoff and Observed Runoff for Validation Data**



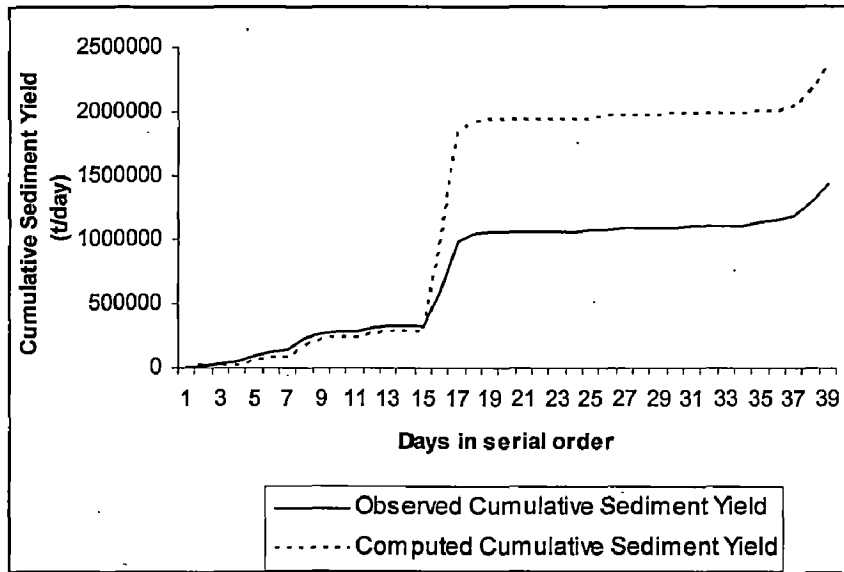
**Fig. 6.3 Scatter Plot of Computed Sediment Yield against Observed Sediment Yield for Validation Data**

### 6.4 Sediment Yield Model

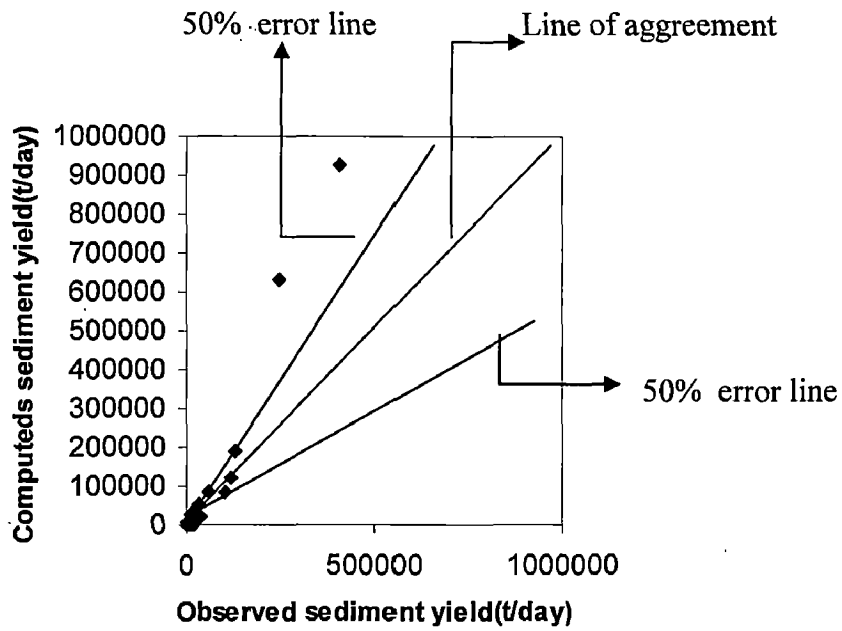
Regression analysis is also performed for sediment yield from the watershed for four sets of input as given in Table 5.3 and their performance are reported in Table 6.2. Sediment yields have been predicted with the 10 percent runoff data as input in the proposed regression models (Table 6.2). The predicted values were compared with the actual measured data. The Fig 6.4 shows the plot of observed and predicted sediment yield for the model 1. Similarly mass curve and scatter plots are presented in Fig 6.5 and Fig 6.6, respectively for the model 1.



**Fig. 6.4 Validation of Computed Sediment Yield with Observed one**



**Fig. 6.5 Mass curve of Computed Sediment Yield and Observed Sediment Yield for Validation Data**



**Fig. 6.6 Scatter Plot of Computed Sediment Yield against Observed Sediment Yield for Validation Data**

## **7.1 Geomorphologic and Vegetational Analysis**

### **7.1.1 Geomorphologic Analysis**

The summary of various morphometric parameters of watershed is presented in Table 7.1 and drainage pattern has been shown in Fig. 4.1. Number of streams of different orders and their corresponding lengths were computed in ARC GIS environment and results are reported in Table 4.1. Analysis of data reveals that the numbers of streams and stream lengths have exponential relationship with corresponding orders. Such variations are shown through Fig.4.2, 4.3 and Eq<sup>n</sup>. (4.1) and (4.2). The Fig. 4.1 indicates that the drainage network of Kankaimai Watershed follows the dendritic pattern. This type of watershed is characterized by homogeneous lithology with less influence of geological structure. This is also supported by bifurcation ratio of 3.9 in this particular case. Bifurcation ratio characteristically ranges between 3.0 and 5.0 for watershed where the influence of geological structure on the drainage network is negligible. The shape parameters like elongation ratio and circulatory ratio (Table 7.1) reveal that the watershed is of almost round shape. Another factor influenced by the shape of basin is form factor. It's value is 0.487 which is on higher side. All these parameter clearly indicate that the water from the different reaches of basin reaches the outlet in less time and cause higher discharge during a shorter period and also there is more possibility of silt load reaching the outlet points.

Relatively higher drainage density of 0.83 km/km<sup>2</sup> for the basins was due to dense drainage network in the mountainous terrain. Relatively higher drainage density of this watershed is also responsible for quick response of sediment yield and runoff. Drainage segment frequency value for the present case is 0.199 that is considered high indicating a relatively high runoff and soil loss. The values of constant channel maintenance of the drainage basin are 1.2 km<sup>2</sup>/km this implies that to maintain one km of stream 1.2 km<sup>2</sup> land area is required.



**Table 7.1 Summary of Geomorphologic Parameters**

S.N.	Geomorphologic Parameters	Value	S.N	Geomorphologic Parameters	Value
1	No of first order streams	185	13	Circulatory Ratio	0.63
2	No of second order streams	38	14	Form Factor	0.487
3	No of third order streams	9	15	Drainage Density	0.83
4	No of fourth order streams	2	16	Channel of Constant maintenance	1.2
5	No of fifth order streams	1	17	Relief (m)	3511
6	Length of first order streams(km)	623.348	18	Relief Ratio	0.0705
7	Length of second order streams (km)	176.637	19	Relative Relief	0.023
8	Length of third order streams(km)	91.844	20	Channel segment frequency	0.199
9	Length of fourth order streams(km)	60.166	21	Total Stream No	235
10	Length of fifth order streams(km)	30.424	22	Total Stream Length (km)	982.42
11	Bifurcation Ratio	3.9	23	Area (km <sup>2</sup> )	1180
12	Elongation Ratio	0.778	24	Perimeter (km)	153.118

The relative relief and relief ratio values are 0.023 and 0.0705, respectively. These values are slightly on lower side but other morphometric parameters play the dominating role for quick response of watershed.

### 7.1.2 Vegetational Analysis

Vegetational analysis revealed that though about 86% of watershed is covered with vegetation, but the land area covered by dense vegetation is comparatively less (Table 4.2). This results in the formation of more number of streams in the areas with thin vegetation compared to the area with dense vegetation and thereby increasing drainage density and channel segment frequency. On the basis of vegetational status the watershed is fairly good. But other morphometric parameters as discussed earlier indicate quick response of watershed to runoff and sediment yield. In monsoon season, during the time of heavy rainfall the high suspended sediment concentration ranging from 2000 ppm to 3000 ppm have been observed.

## 7.2 ANN Models

### 7.2.1 Rainfall-Runoff Model

Rainfall-Runoff model has been developed for Kankaimai watershed using ANN. The daily data of rainfall and runoff from 1995 to 1999 were selected for the learning/calibration and validation of models. The first four and half years data have been selected for learning/calibration and testing and remaining six months data were used for validation. The data used were daily rainfall and temperature; from the three stations (Ilam tea state, Saktim and Kanyam ); and daily runoff data measured at Mainachuli, the outlet of watershed for this particular case.

Eight different ANN models with combination of different input variables were created and their performance was evaluated. The results of models in this study were evaluated by statistical parameters and graphical indicators as discussed in chapter 5. The results in terms of various numerical performance evaluation indicators considered in this study are presented in Table 5.2. Graphical performance evaluation indicators such as linear scale plot of the simulated and observed daily runoff, double mass curve plots of simulated and observed flows and scatter plot of simulated versus observed flows are given in Fig. 5.7 to Fig. 5.9 for model 4.

The Table 5.2 reveals that results produced by ANN *model 1*, which takes the concurrent rainfall values ( $R_{1t}$ ,  $R_{2t}$ ,  $R_{3t}$ ) of three different stations as inputs does not produce very satisfactory results in terms of both graphical and numerical performance indicators. This implies that, for this particular type of watershed the rainfall information alone is not sufficient to compute the runoff from a catchment as the state (related to the soil moisture) of the catchment plays an important role in determining runoff rate behaviour. This was also observed by Minns and Hall (1996) (Rajurkar, 2004).

In *model 2*, in addition to the rainfall data of *model 1* the previous time period runoff ( $Q_{t-1}$ ) has also been considered as one of the inputs. This has produced better results in terms of both graphical and numerical performance indicator as compared to the *model 1*. With the inclusion of additional input  $Q_{t-1}$ , the RMSE value decreased by 35.3  $m^3/s$  and  $R^2$  increased by

0.16. But the correlation coefficient value remained almost same. A larger scatter of data about line of agreement can be observed in *model 1*, whereas relatively smaller scatter has been seen in case of *model 2*. The scattered plot shows that about ninety six percent of data are within 50% error line in case of model 2 where as this figure is 48 percent in model 1 (Table 5.2). Similarly *model 2* gives higher recovery ratio (Table 5.2) i.e., ratio of cumulative predicted runoff to cumulative observed runoff, as compared to *model 1*.

In *model 3*, in addition to the other inputs of *model 2*, daily temperatures at three stations have also been included as inputs considering different lag times as obtained from the cross correlation analysis. Result shows that there is little improvement in  $R^2$ , correlation coefficient, but RMSE is inferior to that of model 2. There is also some improvement in scatter plots and recovery ratio. So from this it becomes obvious that the temperature has some influence on runoff from this watershed. This is also seen in *model 8*, which uses the concurrent rainfall data and temperature data of time period  $t-26$ ,  $t-9$  and  $t-8$  as inputs. With the addition of temperature data as one of the inputs in *model 1* there has been little improvement in the model performance. In this case RMSE,  $R^2$  and coefficient of correlation are  $129.35 \text{ m}^3/\text{s}$ , 0.71 and 0.882, respectively. When this model is compared with *model 1* it can be seen that there has been increase of  $R^2$  by 18.33 percent and decrease of RMSE by 14.36 percent. But coefficient of correlation remains same. The improvement in performance can also be noticed in recovery ratio and scattered plots compared to *model 1*.

*Model 4* has been constituted to see the performance of ANN model with the incorporation of previous day runoff of time step  $t-2$  ( $Q_{t-2}$ ) as additional input in the *model 2*. This combination of inputs in the ANN models performs better than the previous models investigated. The RMSE in the second scenario was  $115.75 \text{ m}^3/\text{sec}$ , which reduced to  $103.67 \text{ m}^3/\text{s}$  in the present case. Similarly there has been improvement in the value of  $R^2$  and CC as can be seen in Table 5.2. The scatter plot (Fig 5.9) shows that there is less scatter of points about agreement line. About ninety four percent data fall under 50 percent error line. The recovery ratio in this case is 0.89 which is higher than that of all models except *model 5*. But *model 5* is inferior to *model 4* in all other respects. Improvement of performance of *model 4* due to inclusion of  $Q_{t-1}$  and  $Q_{t-2}$  reveals that soil moisture condition plays important role in this watershed.

*Model 5* was developed to see the effect of inclusion of antecedent rainfall of time step  $t-1$  in the performance of *model 2*. From the Table 5.2 it can be seen that this scenario is inferior in all aspect of statistical evaluation criterion i.e.,  $RMSE$ ,  $R^2$  and  $CC$ . The graphical indicator and recovery ratio also do not show any improvement in performance of the model. This indicates that the previous day rainfall does not have significant influence on the runoff. This is also supported by the time of concentration which is about eight hours.

*Model 6* includes the previous day's runoffs of time steps  $t-1$  and time step  $t-2$ ; and concurrent rainfall and rainfall of time step  $t-1$  as the inputs. In this scenario all the inputs except  $Q_{t-2}$  are same as in model 5. Similarly antecedent rainfall of time step  $t-1$  has been taken as additional input to the *model 4* to construct this model. Idea of developing this particular model is to investigate the effect of these additional inputs in the model performance. The numerical performance indicators indicate that this model is superior to previous one (*model 5*) in terms of  $RMSE$ ,  $R^2$  and  $CC$ . But still it is not superior to the model 2, 3 and 4. This also reveals that antecedent rainfall does not have any influence to runoff from the watershed.

*Model 7* uses the antecedent runoff of time step  $t-1$  as the input to the ANN model. The graphical and numerical indicators revealed that the previous days runoff alone cannot predict the runoff of the watershed. But in combination with rainfall these runoff data can improve the model performance considerably which is evident from the result of model 2, 3, 4, 5 and 6.

### 7.2.2 Sediment yield prediction Model

ANN models for prediction of sediment yield from Kankaimai watershed have been proposed in the present study. The records of suspended sediment yield which were measured in Kankaimai river at Mainachuli have been adopted for present the study. The sediment yield data of wet seasons are available from the year 2001 to 2003. In total, 387 data were collected from the record. Of the total data, ten percent were used for model validation and rest of the data were utilized for learning/calibration and testing. The input combinations used in this application to estimate suspended sediment values are (i)  $Q_t$ ; (ii)  $Q_t$  and  $Q_{st-1}$ ; (iii)  $Q_t$  and  $Q_{t-1}$ ; and (iv)  $Q_t$ ,  $Q_{t-1}$  and  $Q_{st-1}$ . Where  $Q_t$  and  $Q_{st}$  represent, respectively, the stream flow and sediment yield at day  $t$ .

Like in runoff prediction models the results of models in this study were evaluated by statistical parameters and graphical indicators. The results in terms of various performance evaluation statistics considered in this study are presented in Table 5.3.

The comparison of the observed and estimated suspended sediments is also presented in the form of validation plot, scatter plot and mass curve plot for the *model 1*.

It is seen from the validation plot (Fig. 5.11) that the *model 1* estimates closely follow the observed values, whereas the *model 2* and *3* underestimate the peaks. This is also confirmed by the scatter plots (Fig. 13). The scatter plots show that forty two percent of predicted data fall under the 50% error line in case of *model 1* whereas this figure is on lower sides in *model 2* and *model 3* and *model 4* has the same value as that of *model 1*. Though the recovery ratio in *model 2, 3, 4* are higher than that in *model 1*, the overall performance is not superior as can be seen from Table 5.3.

The *model 1* which consider concurrent runoff as input, is also superior in terms of numerical performance evaluation criteria. The  $R^2$ ,  $RMSE$  and  $CC$  values are 0.933, 19058 ( $t/day$ ) and 0.968 respectively which are quite satisfactory. In rest of the models they are comparatively on lower side though their performance is also equally good.

In the *model 2* where concurrent runoff and antecedent sediment yield have been considered as model input, the  $R^2$  efficiency value and coefficient of correlation value are almost equal or slightly less than that of *model 1*. But in terms of  $RMSE$  and other graphical indicators i.e., scatter plot, this model is inferior to *model 1*.

In the case of the *model 3* and *model 4* the performance parameters, i.e.,  $RMSE$ ,  $R^2$  and graphical indicator i.e. scatter plots clearly reveal that these scenarios of inputs can not improve the model performance.

According to the verification results so far discussed, the models can be ranked as *model 1*, *model 2*, *model 4* and *model 3* in robustness of suspended sediment load estimation. So based on the above discussions it can be concluded that the ANN models with input variable as concurrent runoff can best simulate the sediment yield in the Kankaimai river.

### 7.3 Regression Models

The regression models have been developed for all input scenarios of runoff prediction ANN models and sediment yield prediction ANN models. Nash efficiency and root mean square error (Table 6.1 and 6.2) clearly indicate that the regression approach is not appropriate method for forecasting both runoff and sediment yield. In runoff prediction models maximum Nash coefficient of 0.60 is observed in model 8, in all other models this figure is on lower side. It can be seen in Fig. 6.1 that it underestimates the peak runoff as well as low flows in case of runoff prediction model. Inferiority of models is also seen in scatter plot (Fig. 6.2) and cumulative mass curve plot (Fig. 6.3 5). Similarly the regression model overestimates the peak and underestimates lower sediment load while predicting sediment yield (Fig. 6.4). The scatter plot (Fig. 6.6) and mass curve plot (Fig. 6.5) also do not show any positive result. Nash coefficients in all four sediment yield models are negative. This clearly implies that regression sediment yield model can not be applied in this watershed for forecasting sediment yield.

### 7.4 Comparison of ANN and Regression Models

The runoff prediction model “ $Q_t = f(Q_{t-1}, Q_{t-2}, R_{1t}, R_{2t}, R_{3t})$ ” and sediment yield prediction model”  $Q_{st} = f(Q_t)$ ” have been selected for the comparison. The validation plot of runoff is shown in Fig. 7.1 and that of sediment yield in Fig. 7.2. Similarly cumulative mass curve plot of observed runoff and sediment yield are given in Fig 7.3 and Fig. 7.4, respectively.

The numerical performance indicators (RMSE, R2 and CC) presented in Tables 5.2, 5.3, 6.1 & 6.2 are self explanatory to conclude that ANN models are far superior to convention regression analysis. This is also supported by the mass curve (Fig. 7.3) of runoff. In case of ANN, the mass curve of runoff is close to observed one where as that generated from regression analysis deviated from it. The scatter plot also indicates that the spread in the values computed using ANN model is less than that using regression analysis. The validation plot (Fig. 7.1) shows that the regression models underestimate the peak as well as low flows where as the validation plot of ANN model closely tracks that of observed data.

Fig 7.2 depicts the plot of observed sediment yield and the computed sediment yield, using the ANN and regression model as function of time, in days, marked serially from 1 to 38. It is clearly seen that computed sediment yield using ANN model tracks the observed sediment yield better than that of regression model. The Regression model overestimate the peak and underestimates low sediment load values (Fig 7.4).

Based on the above discussion it can be concluded that the regression model is not capable of capturing complex and non-linear relationship between rainfall and runoff; and runoff and sediment yield where as ANN has demonstrated satisfactory performance in this aspects.

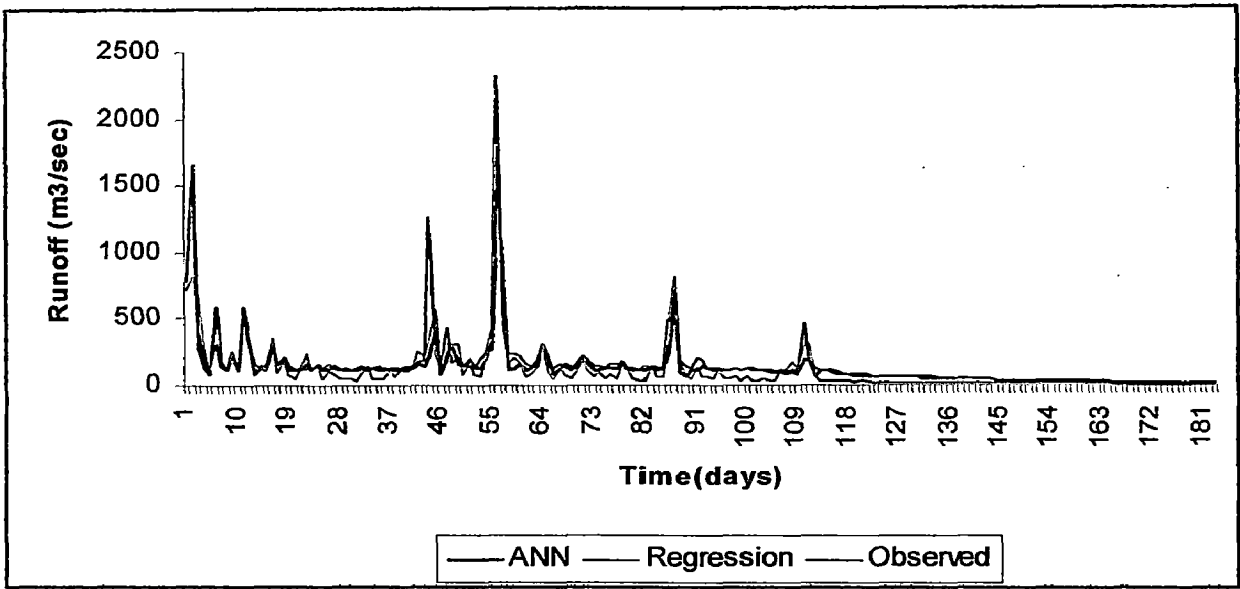


Fig. 7.1 Comparison of ANN and Regression Runoff Prediction Models (Runoff verses Time)

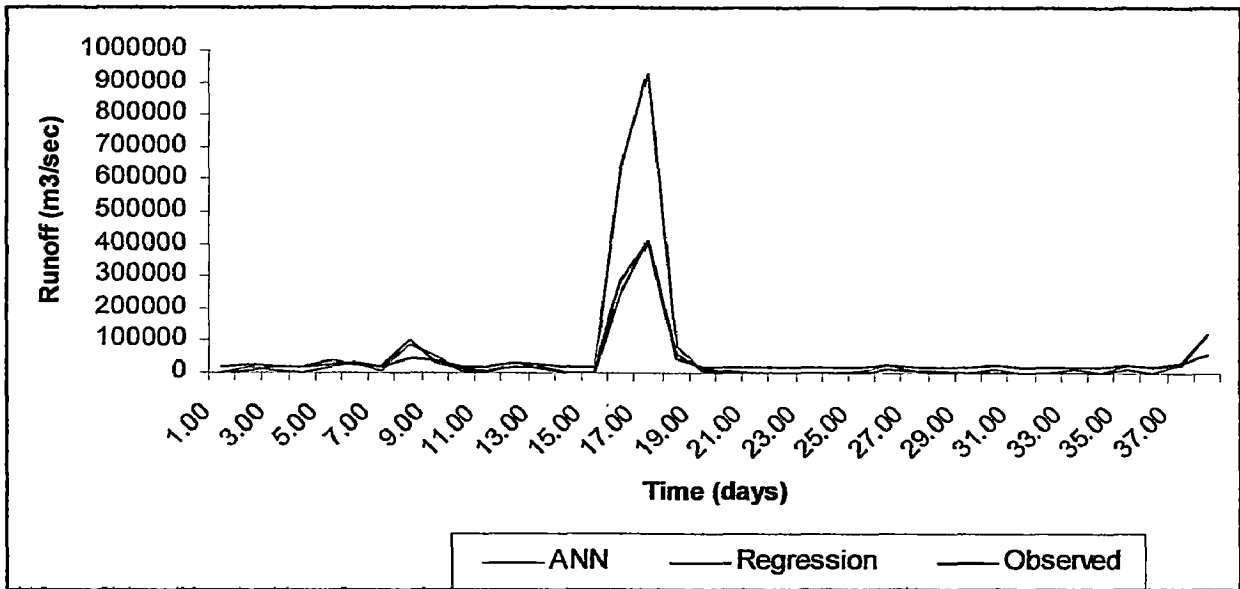


Fig. 7.2 Comparison of ANN and Regression Sediment Yield Prediction Models (Sediment Yield verses Time)



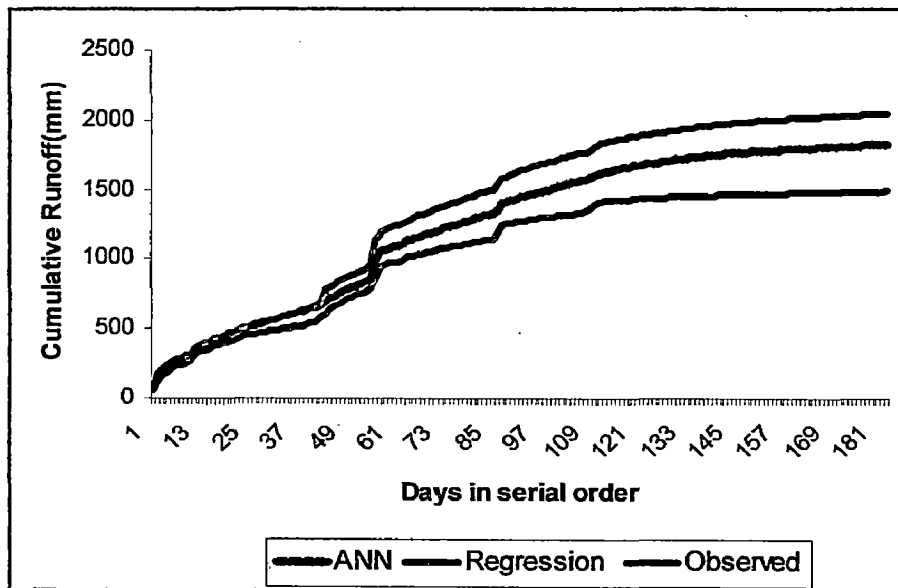


Fig. 7.3 Comparison of ANN Regression Models for Runoff Prediction (Mass Curve)

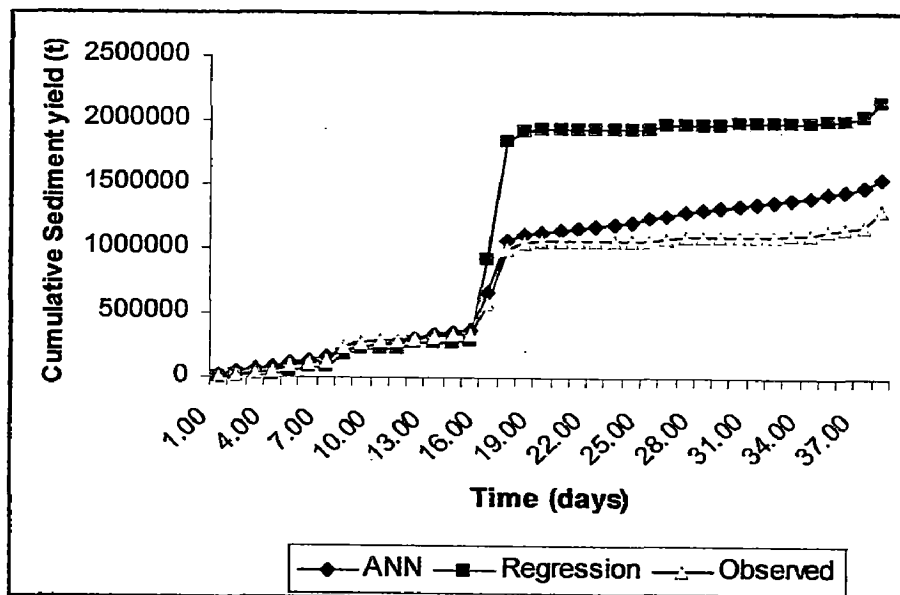


Fig. 7.4 Comparison of ANN Regression Models for Sediment Yield Prediction (Mass Curve)

## CHAPTER –8

### CONCLUSIONS AND RECOMMENDATION

The runoff and sediment yield process is highly non-linear and complex. So it is very important task to develop simulation model to represent these processes. In this study it has been attempted to develop ANNs model for different inputs. All together eight models for runoff prediction and four models for sediment yield were investigated. The results obtained in this study indicate the capability of back propagation ANN models in forecasting daily runoff and sediment yield. Regression models for runoff and sediment yield are also developed to compare their results with ANN models.

Realizing the facts that basin, stream network geometries and vegetation have some sort of relationship with transmission of water and sediment through the basin, the morphometric and vegetational analyses of basin were also carried out in this study.

Based on the results of various investigations following conclusions can be drawn:

1. The Kankaimai watershed is fairly good with moderately high peak flow of shorter duration. The watershed is characterized by homogenous lithology with less influence of geological structure. The drainage density of watershed is high, which promotes quick response of sediment yield and runoff. Number of streams and length of streams have exponential relation with stream order.
2. The basin is predominantly covered by sparse and medium vegetation and have moderately high rate of soil erosion. The land area covered by dense vegetation is comparatively less, which results in the formation of more numbers of streams.
3. For the Kankaimai watershed the rainfall input alone is not sufficient to compute the runoff from the watershed as the state (related to the soil moisture) of the watershed plays an important role in predicting runoff. Inclusion of temperature as input alongwith the rainfall slightly improves the predicting capability of model but overall result is not satisfactory. Performance evaluation of different models suggests that antecedent runoffs of time step 't-1' and 't-2' as an additional inputs variable alongwith rainfall improve the performance of model for this watershed. Antecedent rainfall variables when used as

additional variables can not further improve the predicting capability of models. This implies that for this watershed previous day's rainfalls values have minimal effect on runoff.

4. The concurrent runoff has very high correlation with sediment yield. Use of antecedent runoff and sediment yield as additional input variables to concurrent runoff does not improve the model performance.
5. ANN models are superior to regression models in forecasting runoff and sediment yield in all respects.

## REFERENCES

---

1. Agrawal, A., Singh, R.D., Mishra, S.K., and Bhunya, P.K., (2005). "ANN based sediment yield model for Vamsadhara river basin (India)", ISSN, 0378-4738.
2. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, (2000). "Artificial Neural Network in Hydrology, II Hydrologic Applications", Journal of Engg., ASCE, 115-123.
3. Govindaraju, R.S., Rao, A.R., (1999). "Final Report of ASCE Task Committee on Application of Artificial Neural Networks in Hydrology", ASCE.
4. Jacov, (1993). "Design Guidelines of Irrigation Projects", WECS.
5. Jensen, J.R., (2000). "Remote Sensing of the Environment", Prentice-Hall, 361-362.
6. Jha, S.K., Jain, Ashu, (2005). "Evaluation of ANN technique for rainfall-Runoff Modeling in a large watershed", Proceedings of the international conference on Hydrological perspective for sustainable development -HYPESD, IIT Roorkee 1,180-181.
7. Kisi, O., (2005). "Daily River Flow Forecasting Using Artificial Neural Networks and Auto-Regressive models", Turkish J. Eng. Env. Sci., 29, 9-20.
8. Kisi, O., (2005). "Suspended sediment estimation using neuro-fuzzy and neural network approaches", Hydrological science journal - 50(4), 683-695.
9. Rajurkar, M.P., Kothiyari, U.C., and Chaube, U.C., (2004). "Modeling of the Daily rainfall-runoff relationship with Artificial Neural Networks", J.Hydrology , 285,96-113.
10. Ryan, F., (1999). "Effect of spatial scale on hydrologic modeling", Msc Thesis , Virginia Polytechnic Institute and State University.
11. Sajikumar, N., Thandaveswara, B.S., (1999). "A Non-linear rainfall-runoff using an artificial neural network", Journal of hydrology, 216, 32-55.
12. Shrestha, D.P., (1997). "Assessment of soil erosion in the Nepalese Himalaya, a case study in Likhu Khola valley , middle maountain".Oxford & IBH publishing , vol. 2, no.1, pp. 59-80.
13. Sarangi, A., Madramoofoo, C.A., Enright, P., Prasher, S.O., and Patel, R.M., (2005) "Performance of evaluation of ANN and geomorphology based models for

- runoff and sediment yield prediction for a Canadian watershed”, *Current science*, 89, No 12 ,25.
14. Sen, Z., Altunkaynak,A., (2003). “Fuzzy awakening in rainfall-runoff modeling”, *Nardic hydrology*, IWA publishing. Vol 35 No 1, pp 31-45.
  15. Singh,V.P., (1989).“Hydrologic System. Vol.II Watershed Modelling”, Prentice-Hall, Englewood Clifs, New Jersey, 07632 107-136.
  16. Singh,V.P., Woolhiser, D.A., (2002).“Mathematical Modeling of Watershed hydrology”, *Journal of hydrologic engineering*, ASCE, 4(270) ,270-285.
  17. Subramanya, K., (1994). ” *Engineering Hydrology* “, Second Edition, Tata McGraw-Hill Publishing Company Ltd 151-152, 235.
  18. Verstappen, H.Th., (1983). “Applied Geomorphology”, *Geographical survey for environmental development* ,Elsevier Ltd. 57-72.
  19. Yazidi, B., (2003) “ A Comparative study of soil erosion modelling in Lomkao-Phetchabun, Thailand” , Msc Thesis, ITC, The Netherlands.

## CROSS REFERENCES

---

1. Bedient, P.B., and Huber, W.C.,(1992). "Hydrology and Floodplain Analysis", 2<sup>nd</sup> edition. New York: Addison-Wesley Publishing Company.
2. Campolo, M., Andreussi, P., and Soldati,A., (1999). " River flood forecasting with neural network model", Water Resources Research , 35(4), 1191-1197.
3. Dawson, C.W., and Wilby, R.,(1998). "An ANN Approach to Rainfall-Runoff Modeling", Hydrological Science Journal, 43(1), 47-66.
4. Duan, D., Fermini, B., and Natteli, S., (1971) " Sustained Outward Current Observed after Itol Inactivation in Rabbit Atrial Myocytes Is a Novel CI", Am. J. Physiol, 263: H1967, 992.
5. Foster, G.R., (1982). "Modelling the erosional process", p. 297-380, *In* Haan, ed. Hydrological modeling of small watersheds, 5 ed. ASAE monograph.
6. Halff, A.H., Halff, H.M., and Azmoodeh, M., (1993). "Predicting Runoff from Rainfall Using Neural Network". Proc. Engg. Hydrol., ASCE, NY, 760-765.
7. Hjelmfelt, A.T., and M. Wang, (1993). "Artificial neural network as unit hydrograph application", Engineering Hydrology, Symposium sponsored by hydraulics Division/ ASCE, July 25-30, 1993, San Francisco, Calif.
8. Hsu, K.L., Gupta, H.V., and Sorooshian, S., (2001). "River Flow Time Series Prediction with a Range Dependant Neural Network", Hydrol. Sci.J., 46(5), 729-745.
9. Imrie, C.E., Durucan, S., and Korre, A., (2000). "River Flow Prediction using Artificial Neural Networks: Generalization Beyond the Calibration range", J. Hydrology, 233, 138-153.
10. Jaroslav, M.H., Marcel, R.S., Grešáka, J.B., and Geografický, S.B., (1996). "Modelling spatial and temporal changes of soil water erosion", Geografický casopis 48:255-269.
11. Karunanithi, N., Grenney, W.J., Whitely, D. and Bovee, K., (1994). "Neural Networks for River Flow Prediction ", J. Comp. Civil Eng., Asce8, 201-220.
12. Keneth, G., G. Foster, Glenn, A., and Jeffery, P., (1991). Revised Universal Soil Loss Equation. Journal of soil and water conservation Volume 46(1).

13. Kilgore, J., (1997). "Development and Evaluation of a GIS-Based Spatially Distributed Unit Hydrograph Model", MS BSE Thesis.
14. Kothari, U.C., (1995). "Estimation of Monthly Runoff from Small Catchments in India", Journal of Hydrologic Sciences, Proc. International Association for Scientific Hydrologic (IAASH), 40(4), 535-541.
15. Kuiper, E., (1971). "Water Resources Development", 3<sup>rd</sup> ED. Butterworths, London.
16. Lal, R., (2001). "Soil degradation by erosion. Land degradation and development", 12:519-539.
17. Mamdani, E.H., (1974). "Application of fuzzy algorithms for simple dynamic plant. Proc", IEE, 121, 1585–1588.
18. Morgan, R.P.C., Morgan, D.D.V. and Finney, H.J., (1984). "A predictive model for the assessment of erosion risk", J. Agricultural Engineering Research 30: 245 – 253.
17. Mulvaney, T.J., (1851). "On the use of self-registering rain and flood gauges in making Observations of the relations of rainfall and of flood discharges in a given catchment", Transactions of the Institution of Civil Engineers, Ireland, 4(2): 18.82.
18. Pappis, C.P. and Mamdani, E.H., (1977). "A fuzzy controller for a traffic junction", IEEE Trans. Syst. Ma Cybern., 7(10), 707–717.
19. Renard, K.G., Foster, G.R., Weesies, G.A., McCool, D.K. and Yoder. D.C., (1997). "Predicting Soil Erosion by Water: A guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE)", Agriculture Handbook 703 USDA, Washington.
20. Renschler, C., (19996). "Soil erosion risk mapping by means of geographical information systems (GIS) and hydrologic modeling", MSc, Technical University of raunschweig, raunschweig.
21. Roo, A.P.J., (1993). "Modelling surface runoff and soil erosion in catchments using Geographical Information Systems", PhD, Utrecht University, Utrecht.
22. Schumm, S.A., (1964). "Airphotos and waree Resources", Trans. UNESCO Symp. Aerial Surveys and Integrated Studies, Tolouse, UNESCO: 70-80.
23. S, en, Z. (1998). Fuzzy algorithm for estimation of solar irradiation from sunshine duration. Solar Energy, 63(1),39–49.
24. S, en, Z. (2001). Bulanýk mantýk ve modelleme ilkeleri, Bilgi Ku"ltu" r Sanat Publication House (Fuzzy Logic and Modeling principles book in Turkish), 176pp.
25. Shamseldin, A.Y., (1997). " Application of a Neural Network Technique to Rainfall-Runoff Modeling", J. Hydrol., 199, 272-294.

26. Sherman, L. K., (1932). "Stream flow from rainfall by unit-graph method. Engineering News Record, 108: 501-505.
27. Smith, J., and Eli, R.N., (1995). "Neural Network Models of Rainfall-runoff Process". J. Water Resources Plng. Mgnt., ASCE, 121(6), 499-508.
28. Sorooshian, S., Daun, Q. and Gupta, V.K., (1993). "Calibration of Rainfall-Runoff model: Application of Global Optimization to the Sacramento Soil Moisture Accounting Model", Water Resour. Res., 29, 1185-1194.
29. Sorooshian, S. and Gupta, V.K., (1995). "Model Calibration. In Computer Models of Watershed Hydrology", ed. Vijay P. Singh. Colorado: Water Resources Publications, pp. 23-68.
30. Sugeno, M. (Ed.), (1985). Industrial Applications of Fuzzy Control, North-Holland, New York.
31. Todini, E., (1998). "Rainfall-runoff modelling—past, present, and future", Journal of Hydrology, 100: 341- 352,
32. Tokar, A.S., Johnson, P.A., (1999). "Rainfall-runoff modeling using artificial neural networks", Journal of Hydrologic Engineering, ASCE 4, 232-239.
33. Yapo, P., Gupta, V.K., and Sorooshian, S., (1999). "Calibration of Conceptual Rainfall-Runoff Models: Sensitivity to Calibration Data", J. Hydrol., 181, 23-48,
34. Zadeh, L.A., (1965). "Fuzzy sets, information and control", 8, 338-353.
35. Zealand, C.M., Burn, D.H. and Simonovic, S.P., (1999). "Short-Term Stream flow forecasting Using Artificial Neural Networks", J.Hydrol., 214-, 32-48.



# APPENDIX – A DAILY TEMPERATURE FROM 1995 TO 1999

---

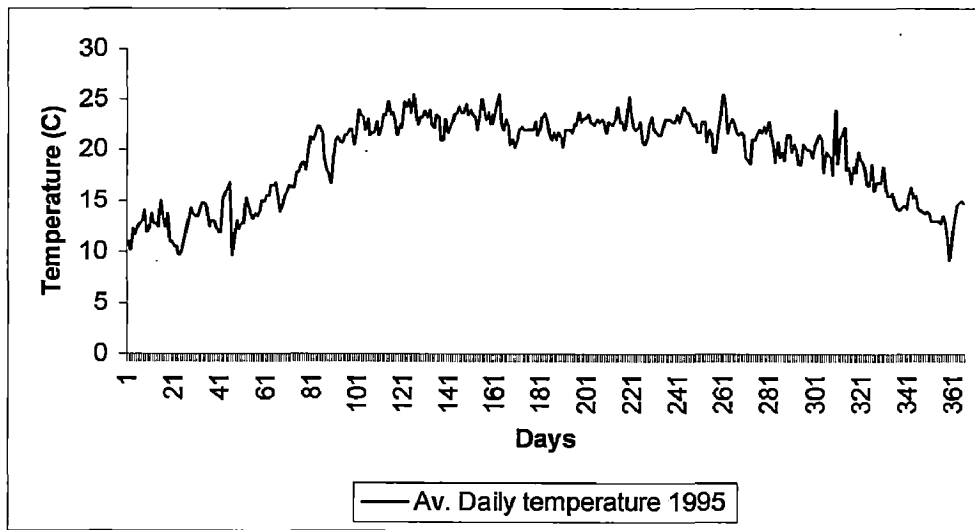


Fig. A1 Daily temperature data recorded at Ilam tea station

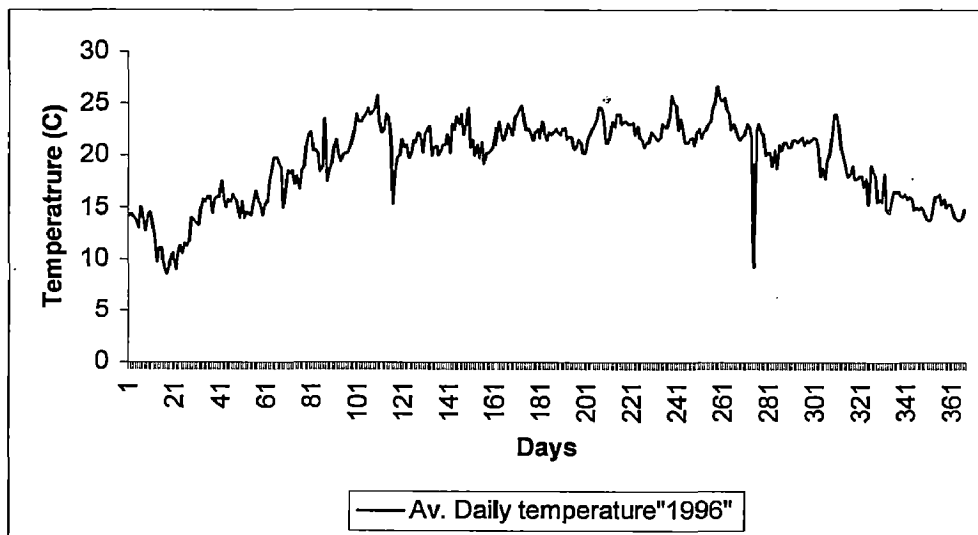
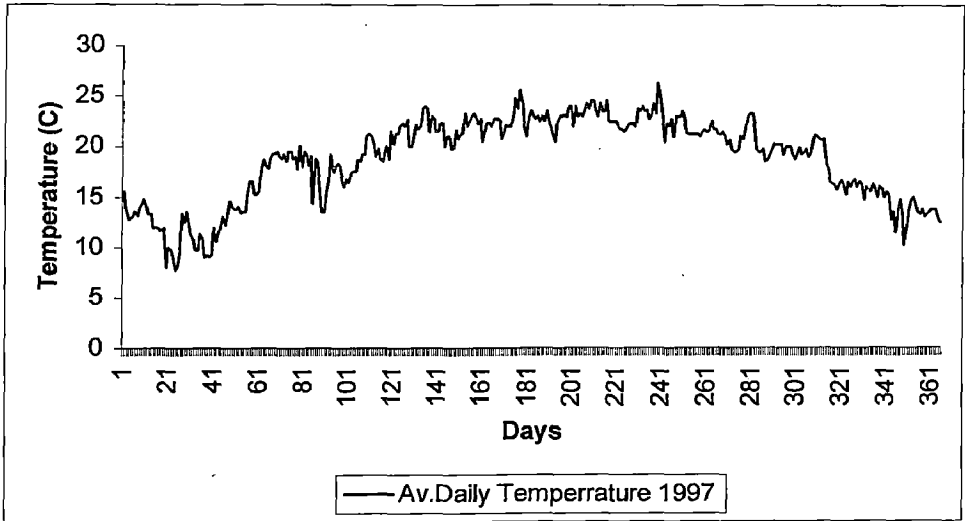
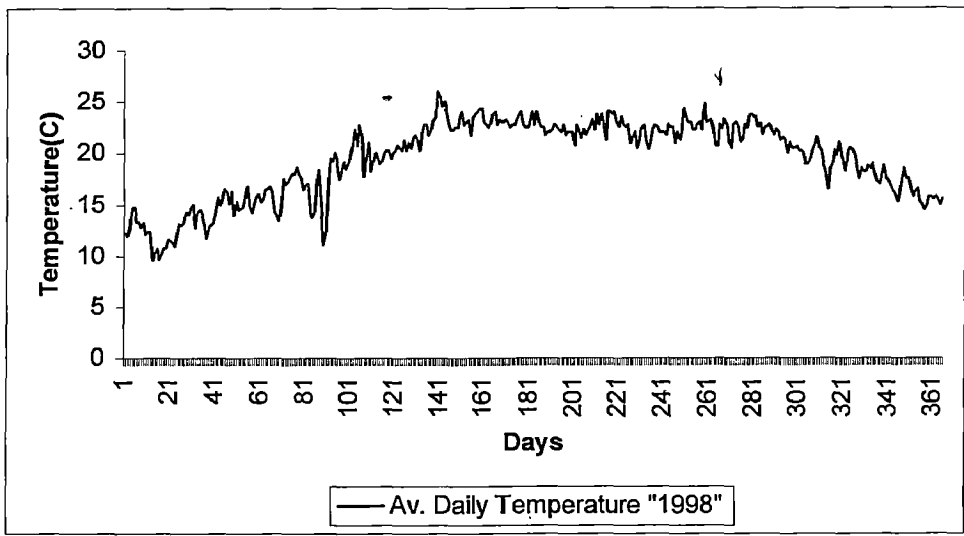


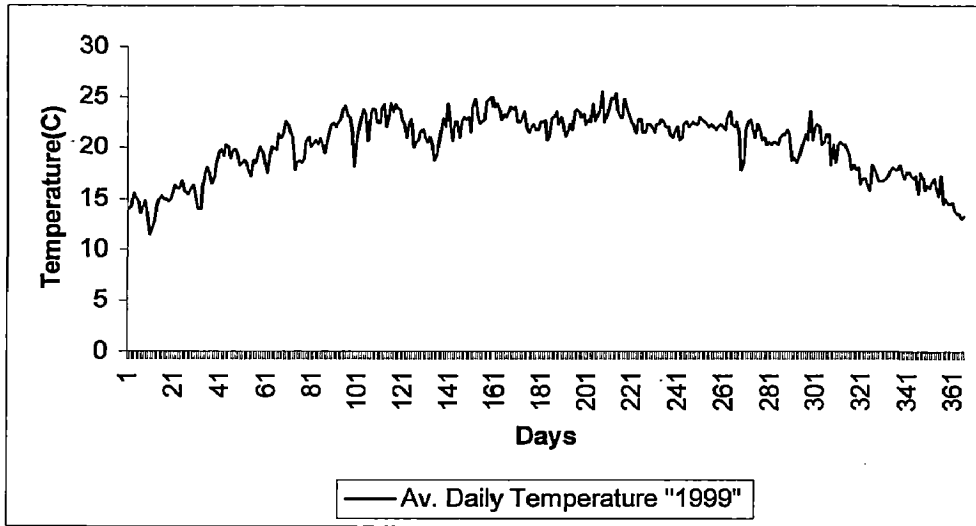
Fig. A2 Daily temperature data recorded at Ilam tea station



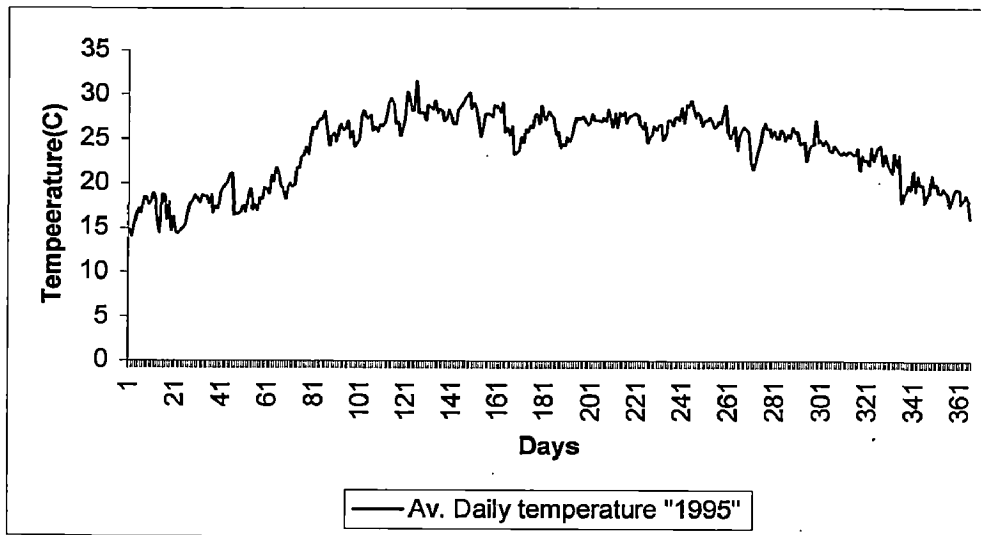
**Fig.A3 Daily temperature data recorded at Ilam Tea Station**



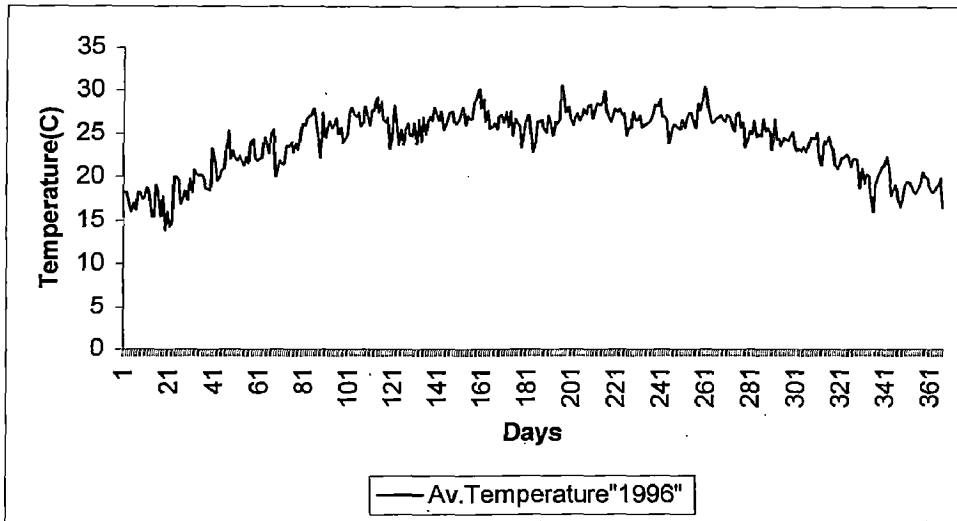
**Fig. A4 Daily temperature data recorded at Ilam tea station**



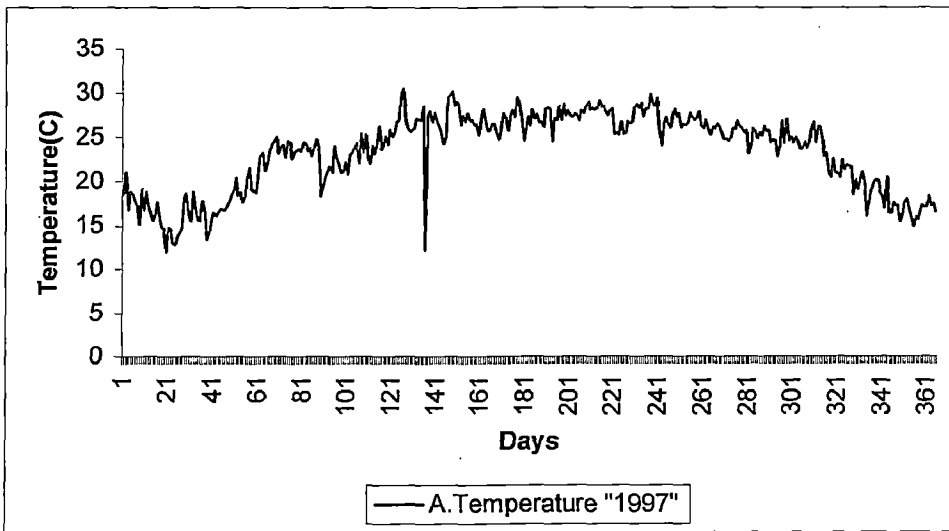
**Fig. A5 Daily temperature data recorded at Ilam tea station**



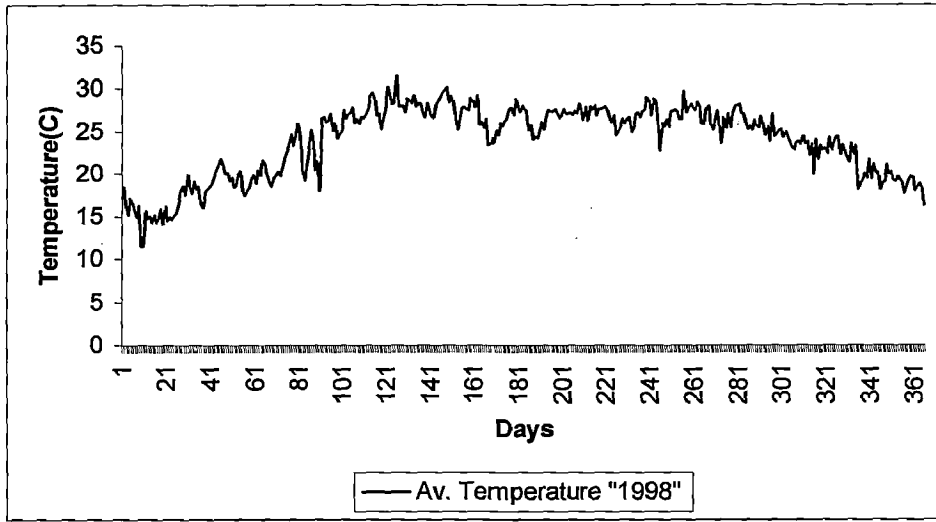
**Fig.A6 Daily temperature data recorded at Saktim station**



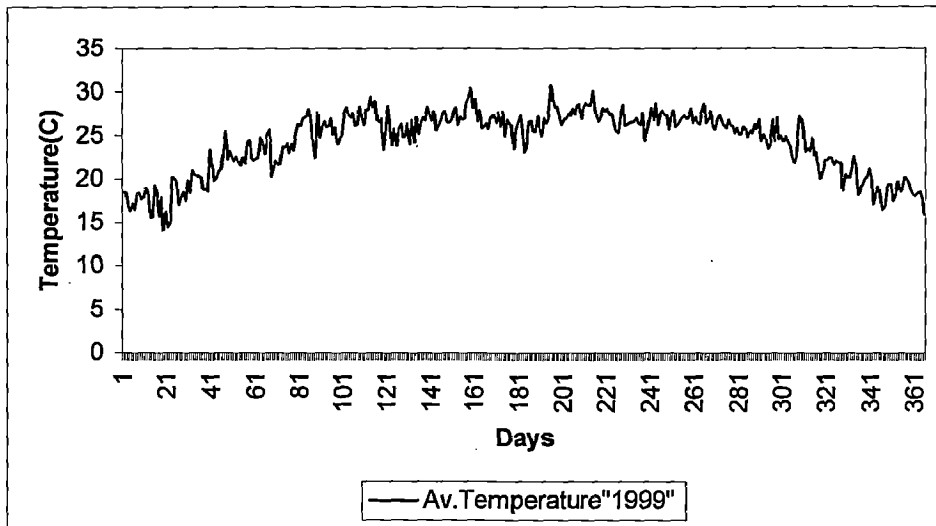
**Fig. A7 Daily temperature data recorded at Suktim station**



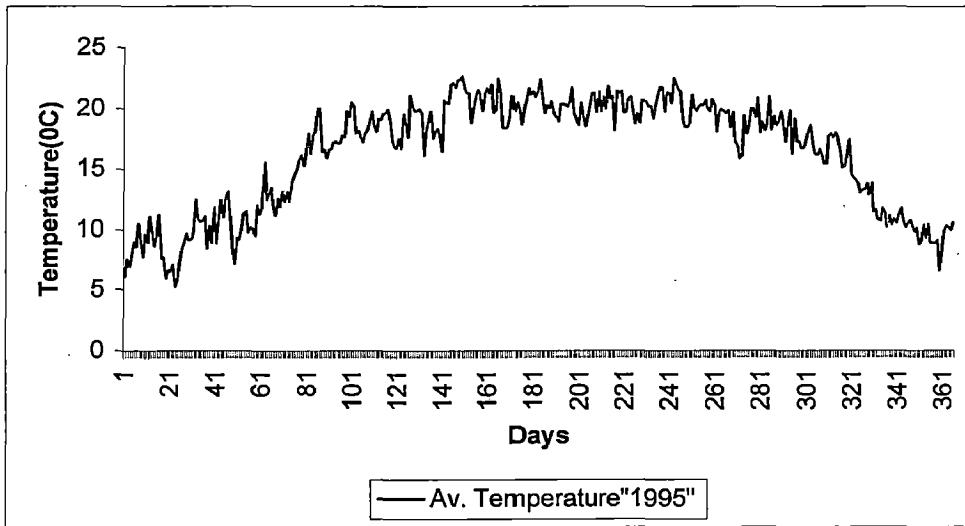
**Fig. A8 Daily temperature data recorded at Suktim station**



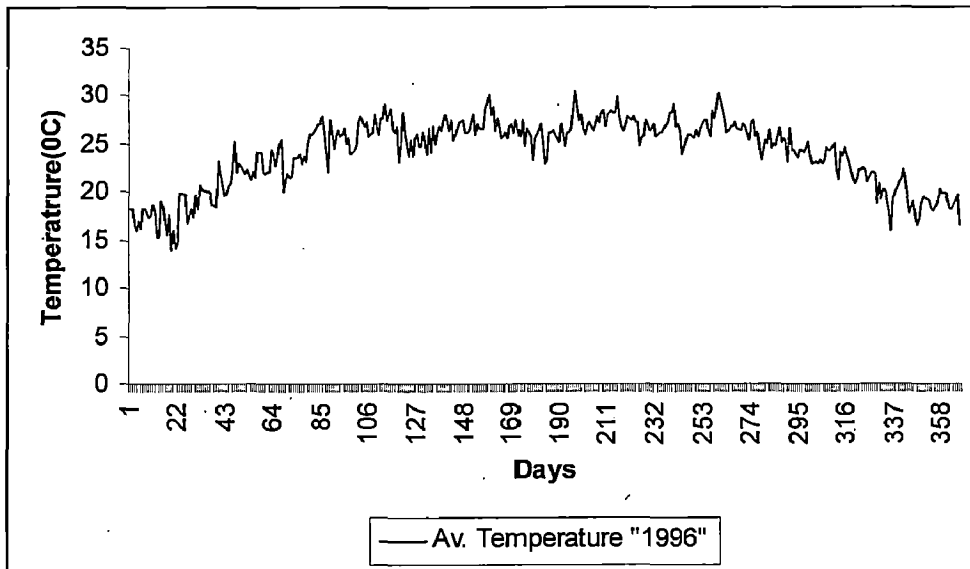
**Fig. A9 Daily temperature data recorded at Saktim station**



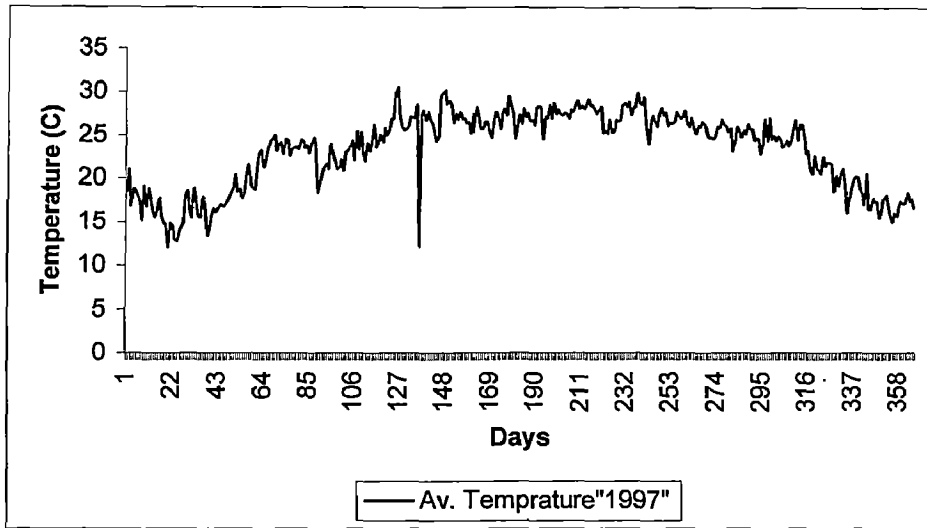
**Fig. A10 Daily temperature data recorded at Saktim station**



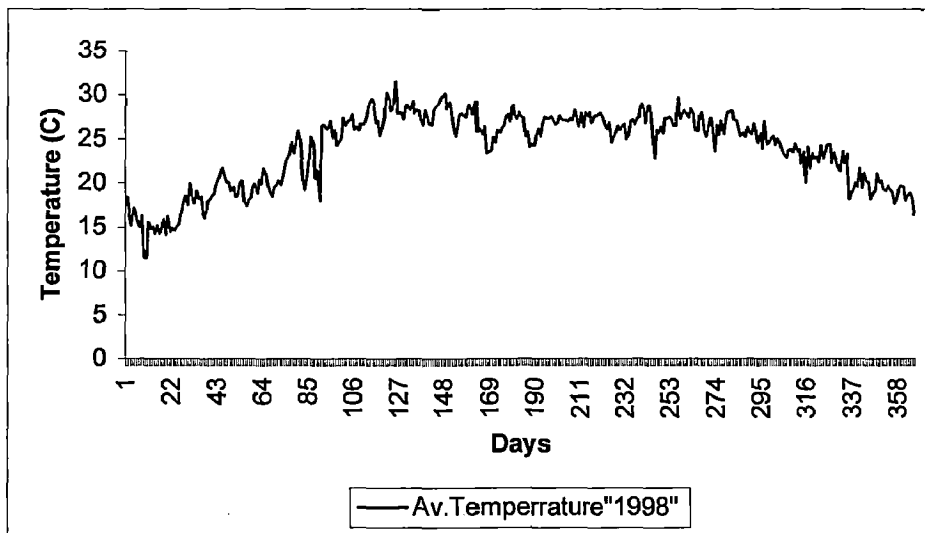
**Fig.A11 Daily temperature data recorded at Kanyam station**



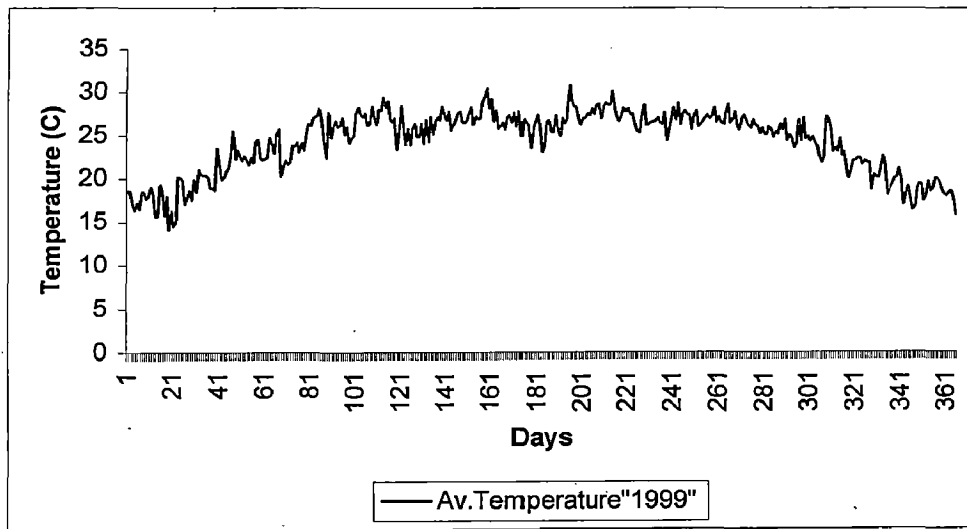
**Fig. A12 Daily temperature data recorded at Kanyam station**



**Fig.A13 Daily temperature data recorded at Kanyam station**



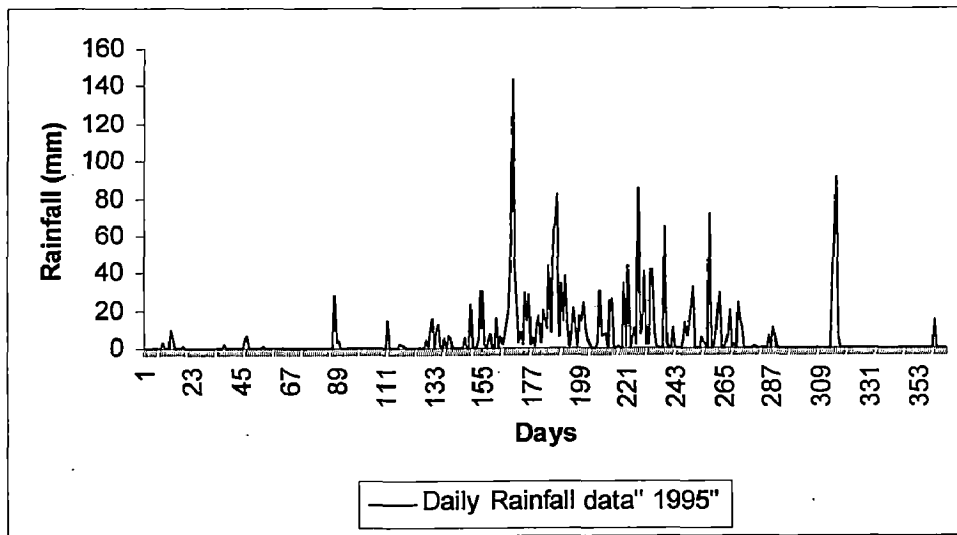
**Fig. A14 Daily temperature data recorded at Kanyam station**



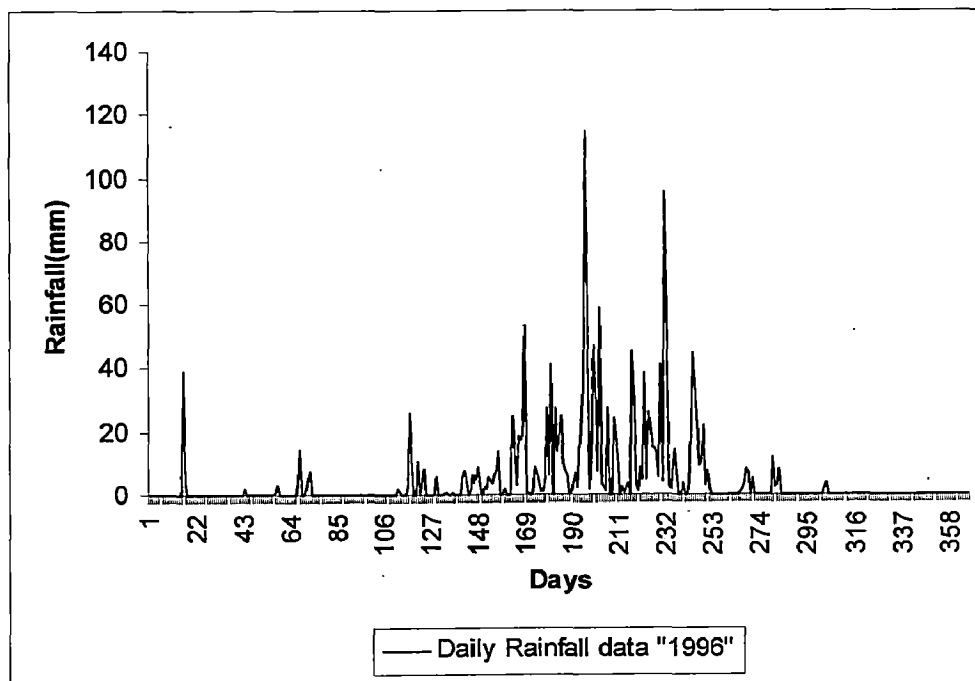
**Fig. A15 Daily temperature data recorded at Kanyam station**



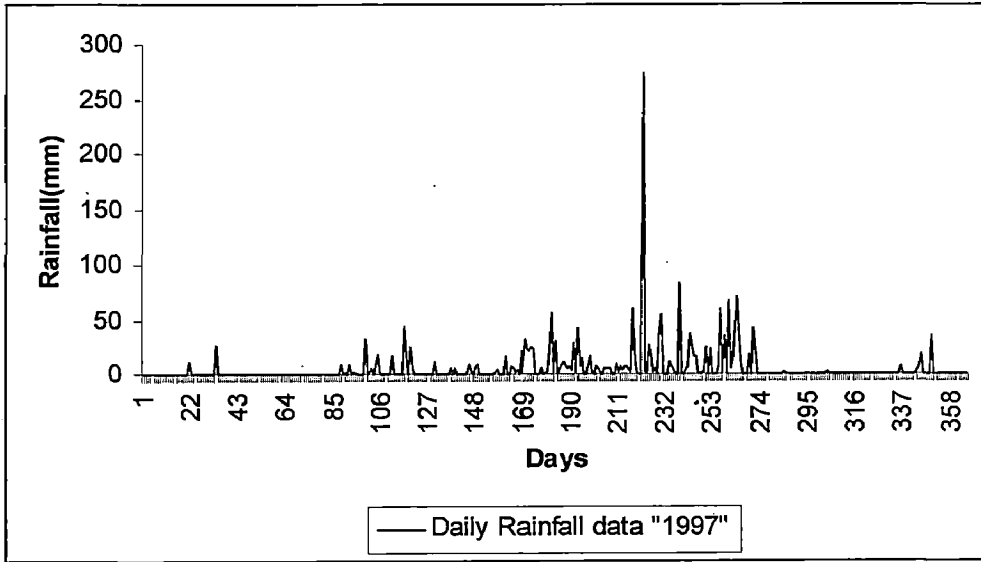
## APPENDIX – B DAILY RAINFALL FROM 1995 TO 1999



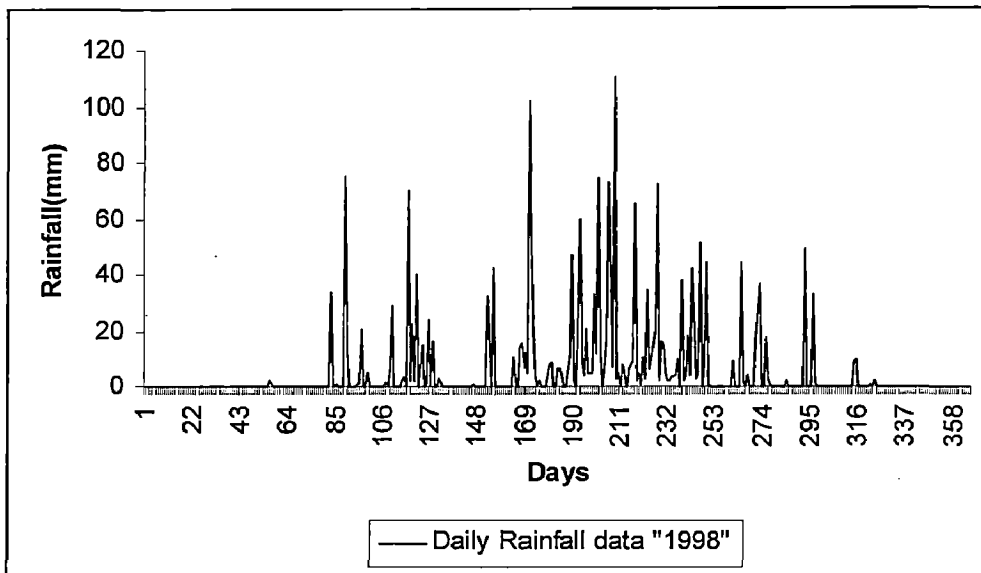
**Fig. B1** Daily rainfall data recorded at Ilam tea station



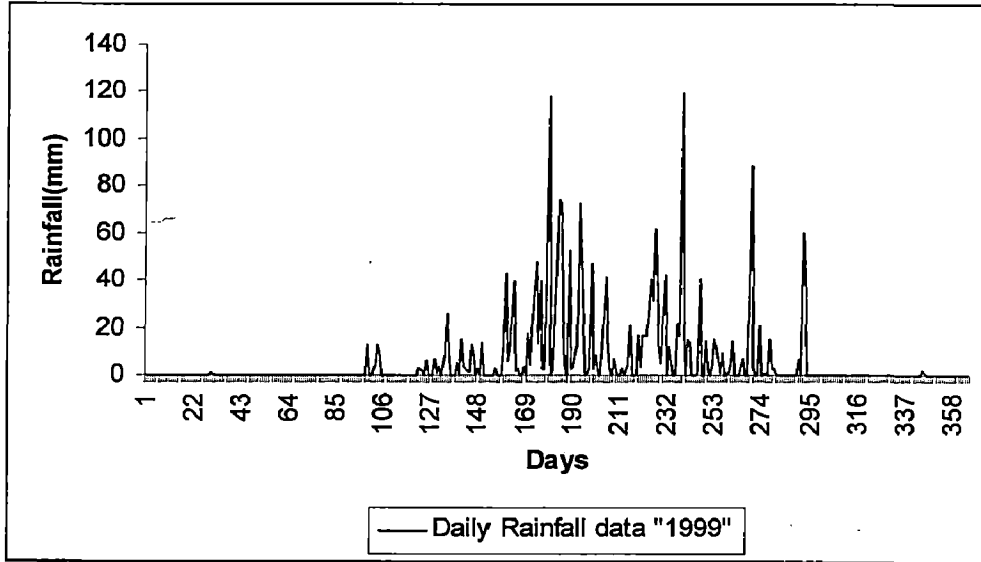
**Fig. B2** Daily rainfall data recorded at Ilam tea station



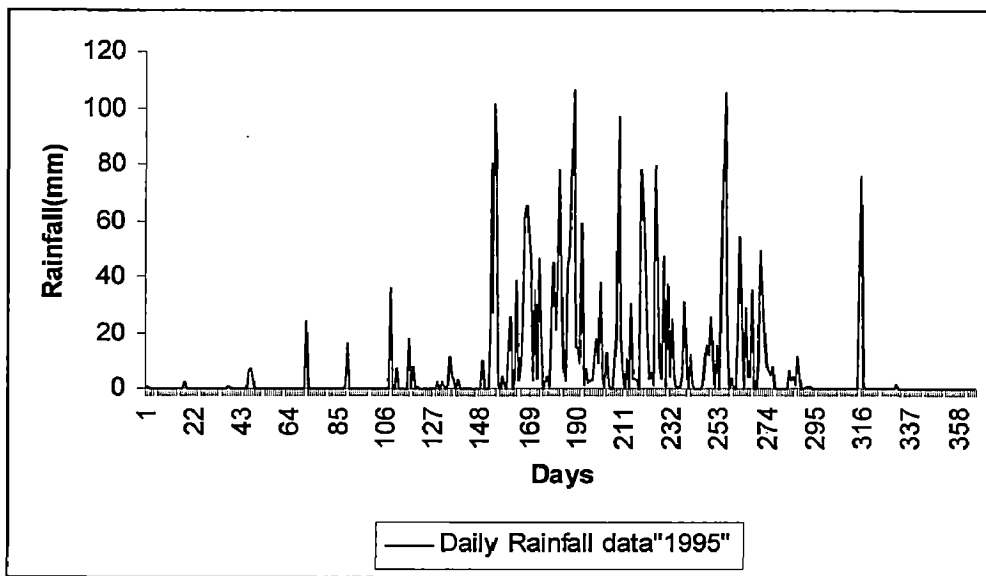
**Fig. B3 Daily rainfall data recorded at Ilam tea station**



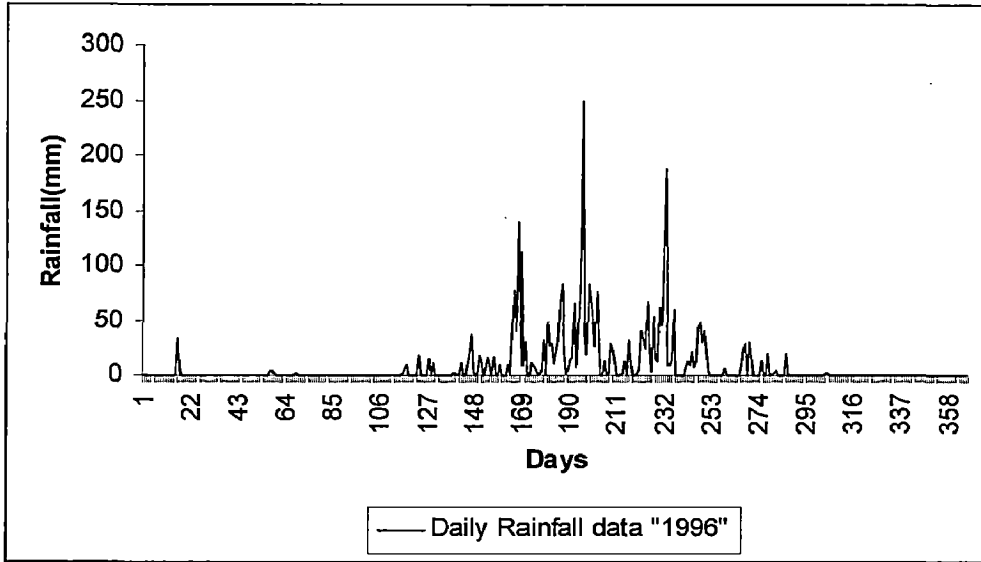
**Fig. B4 Daily rainfall data recorded at Ilam tea station**



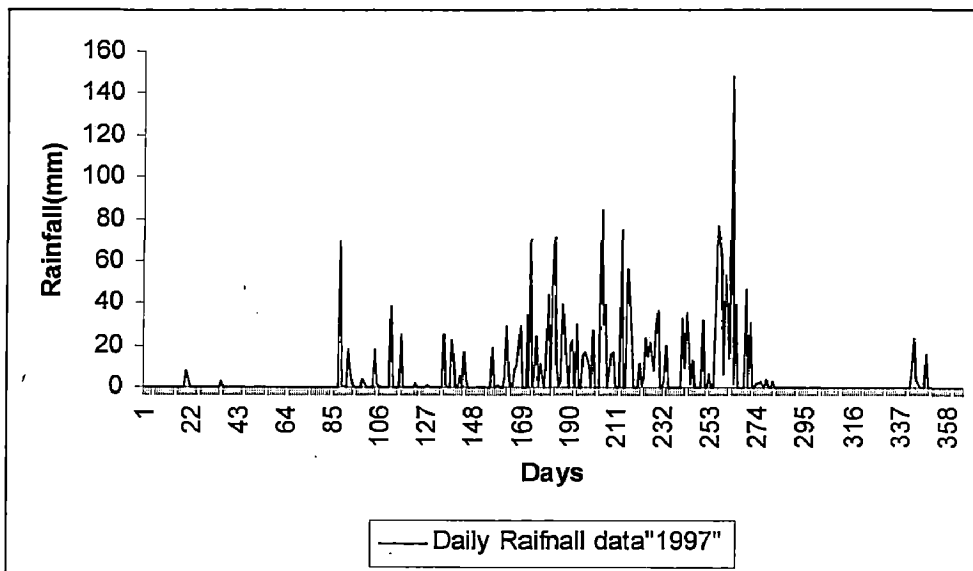
**Fig. B5 Daily rainfall data recorded at Ilam tea station**



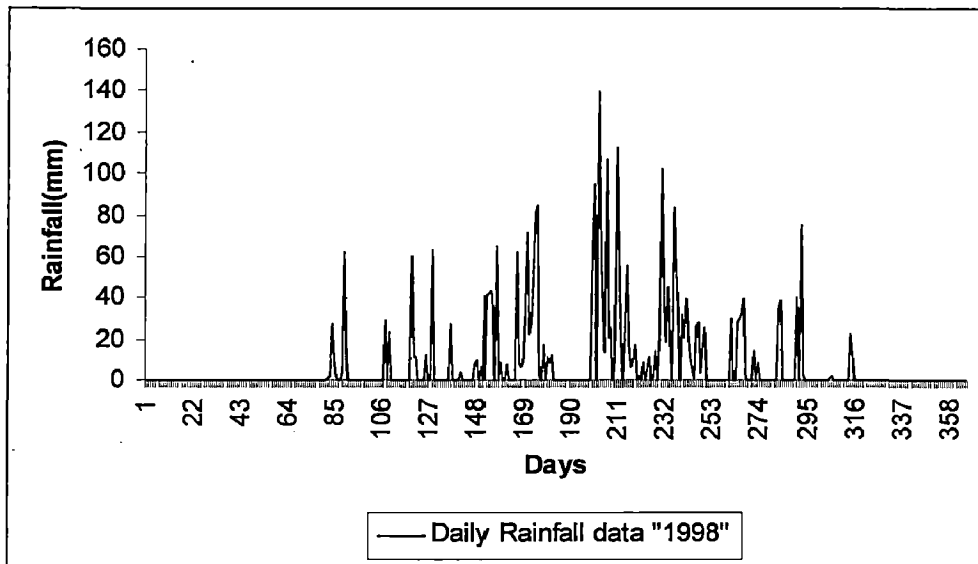
**Fig. B6 Daily rainfall data recorded at Saktim station**



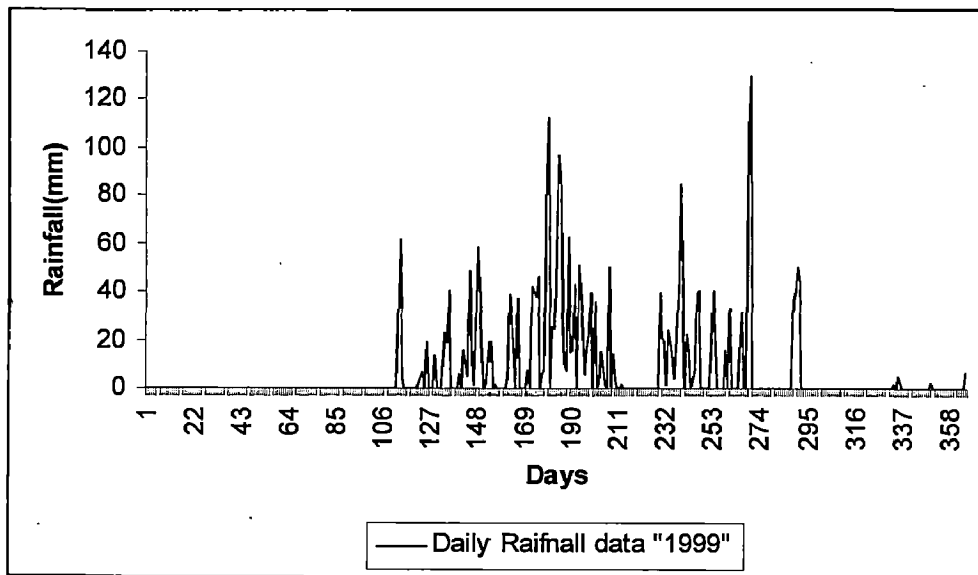
**Fig. B7 Daily rainfall data at Saktim Station**



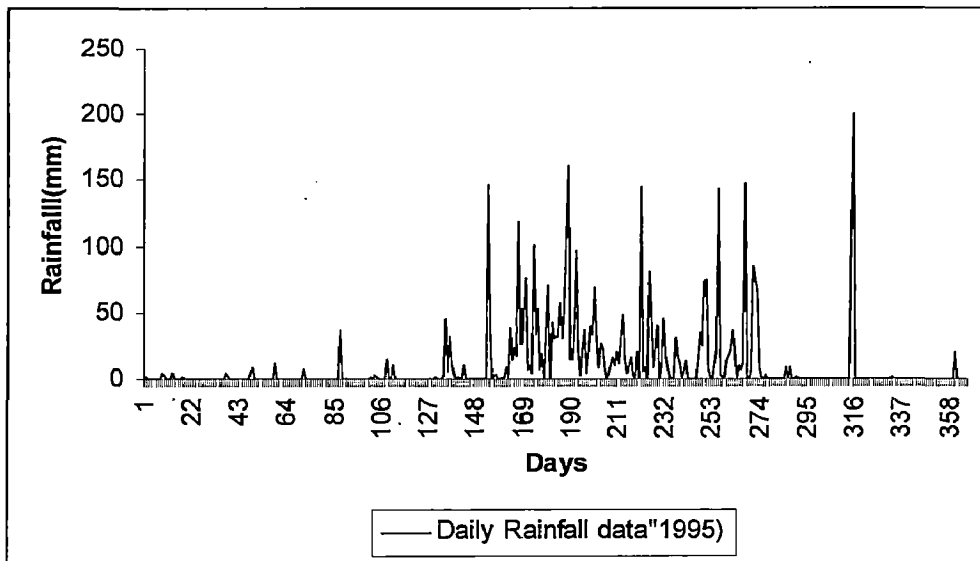
**Fig. B8 Daily rainfall data recorded Saktim Station**



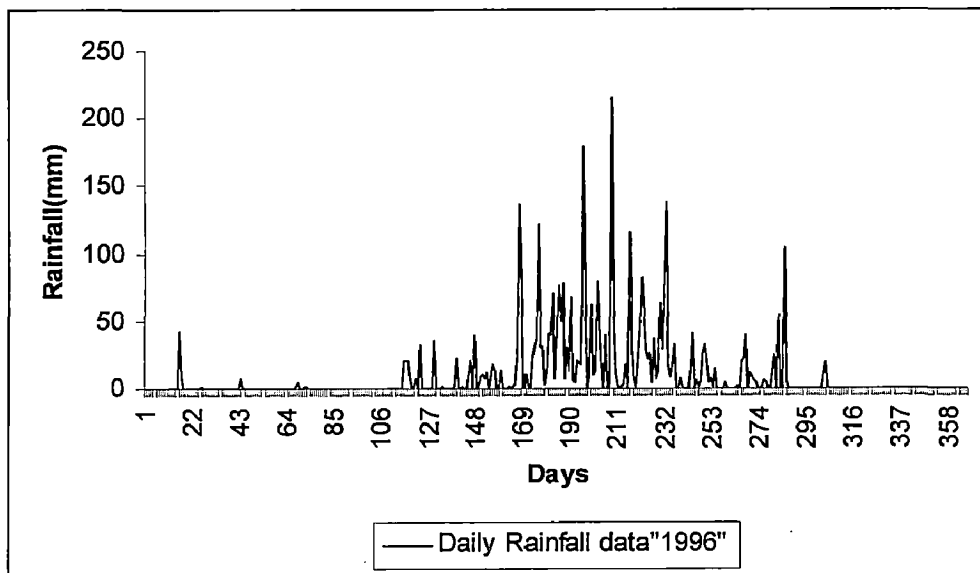
**Fig. B9** Daily rainfall data recorded at Saktim Station



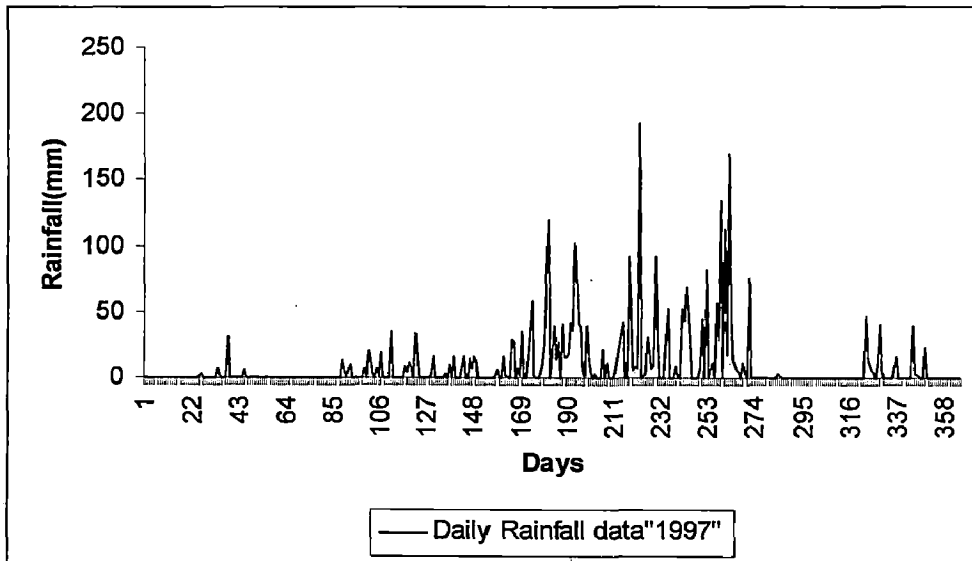
**Fig. B10** Daily rainfall data recorded at Saktim Station



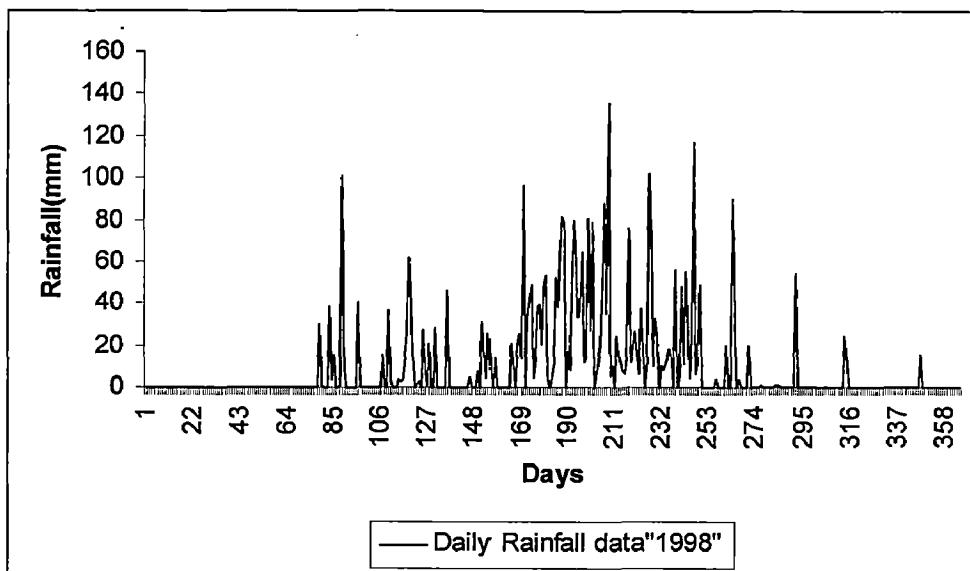
**Fig. B11 Daily rainfall data recorded at Kanyam Station**



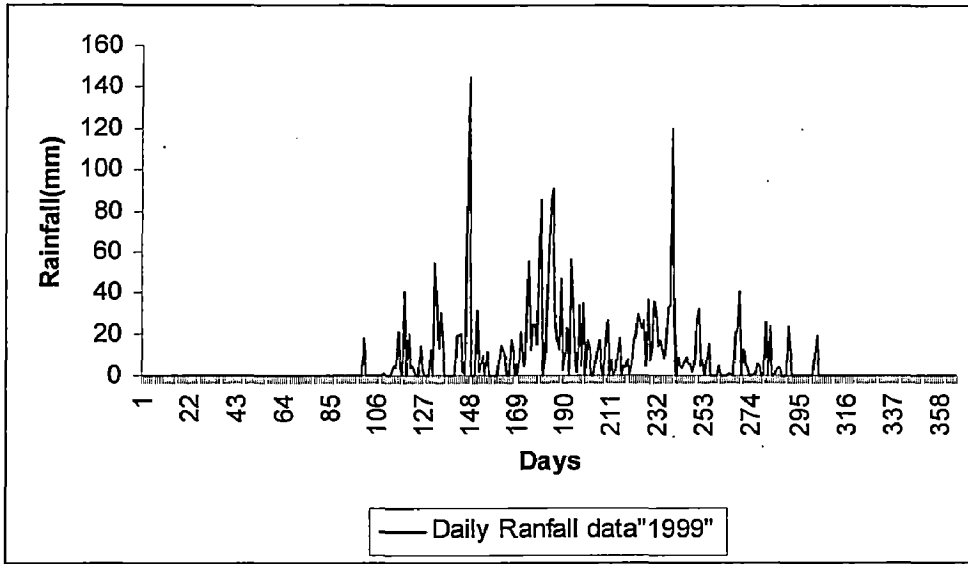
**Fig. B12 Daily rainfall data at Kanyam Station**



**Fig.B13 Daily rainfall data recorded at Kanyam Station**



**Fig. B14 Daily rainfall data at Kanyam Station**



**Fig.B15 Daily rainfall data at Kanyam Station**



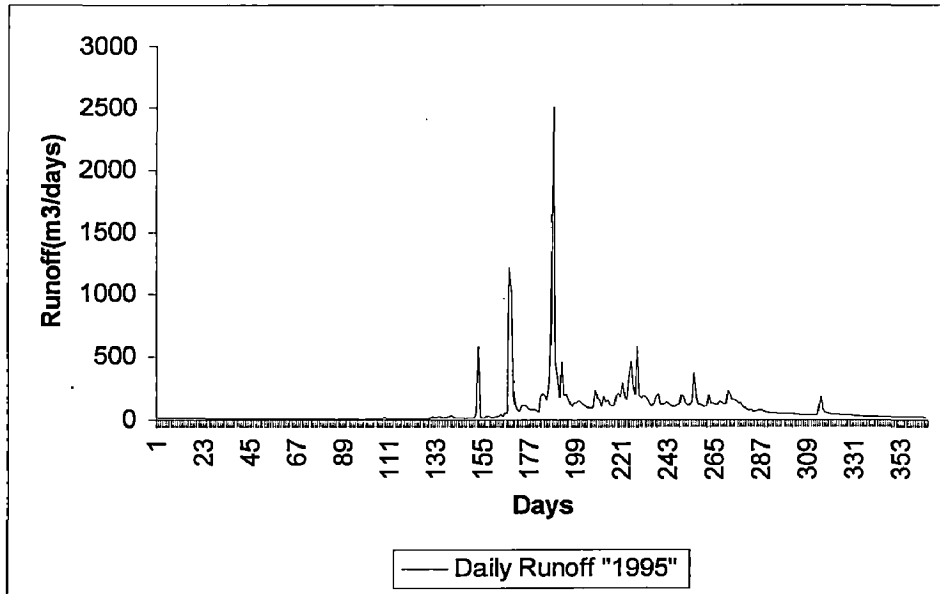


Fig. C1 Daily runoff recorded at Mainachuli

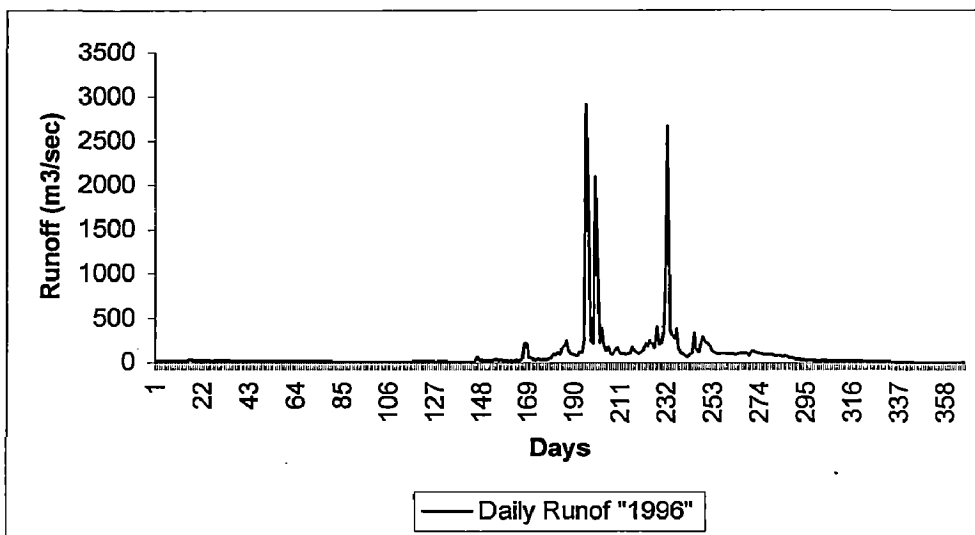
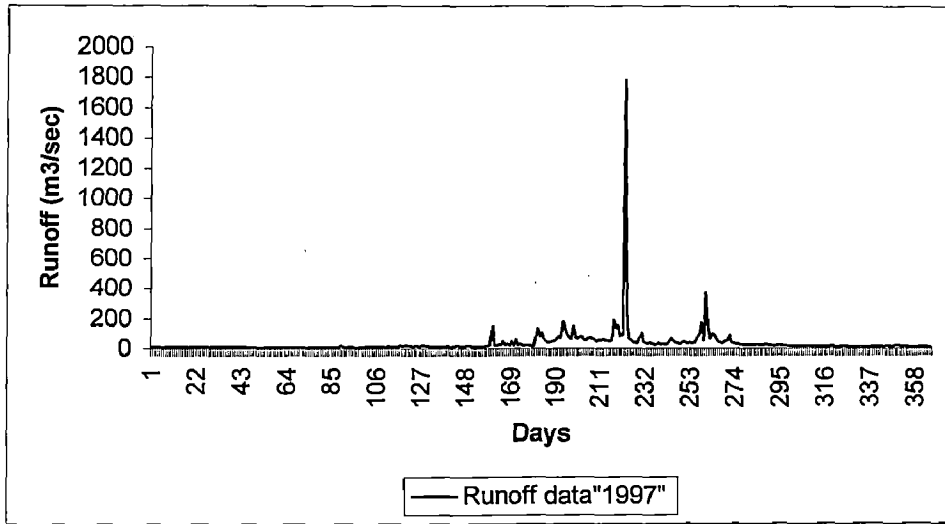
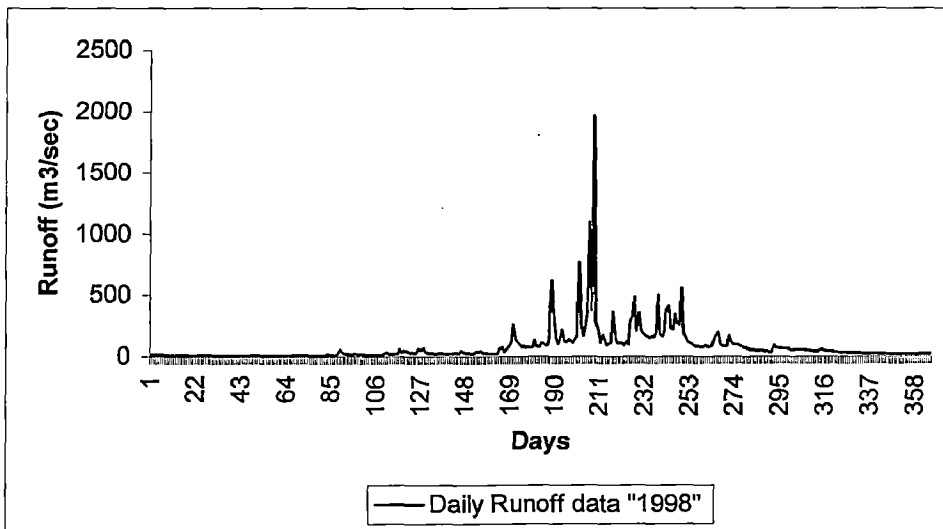


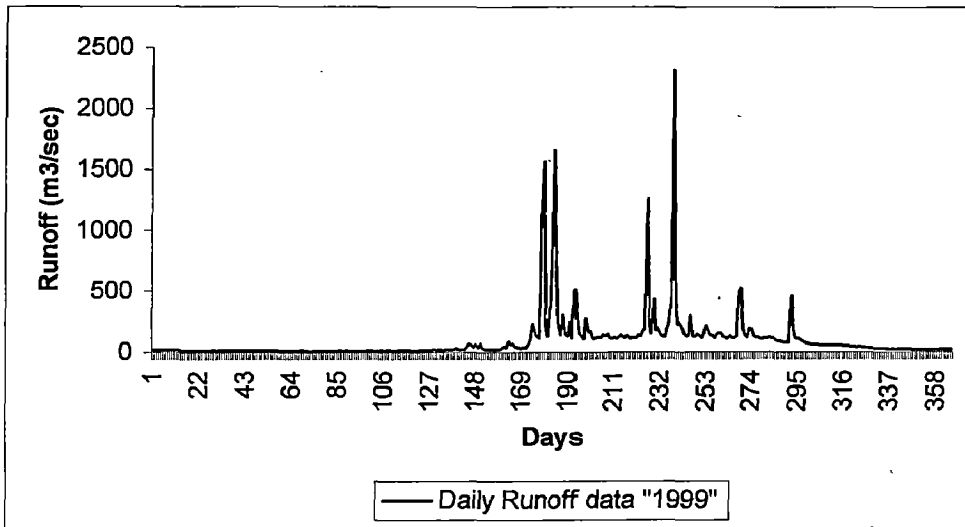
Fig. C2 Daily runoff recorded at Mainachuli



**Fig. C3 Daily runoff recorded at Mainachuli**



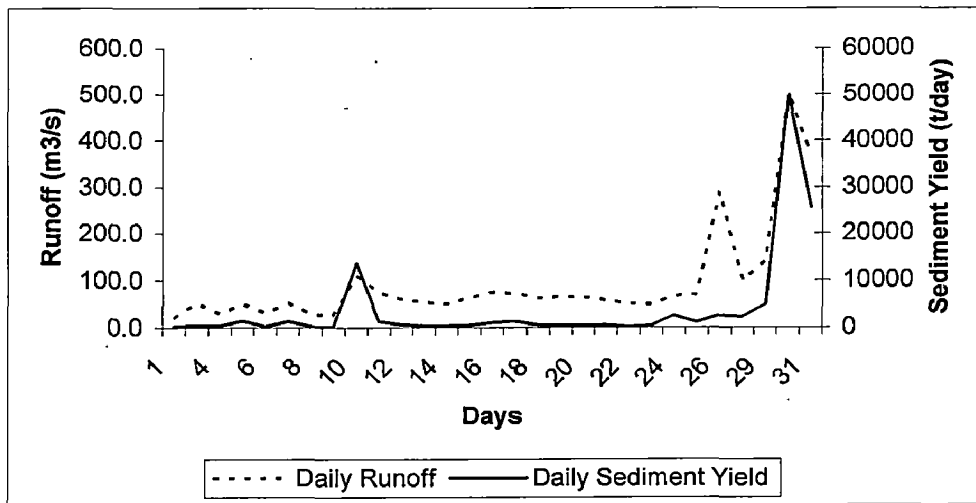
**Fig. C4 Daily runoff recorded at Mainachuli**



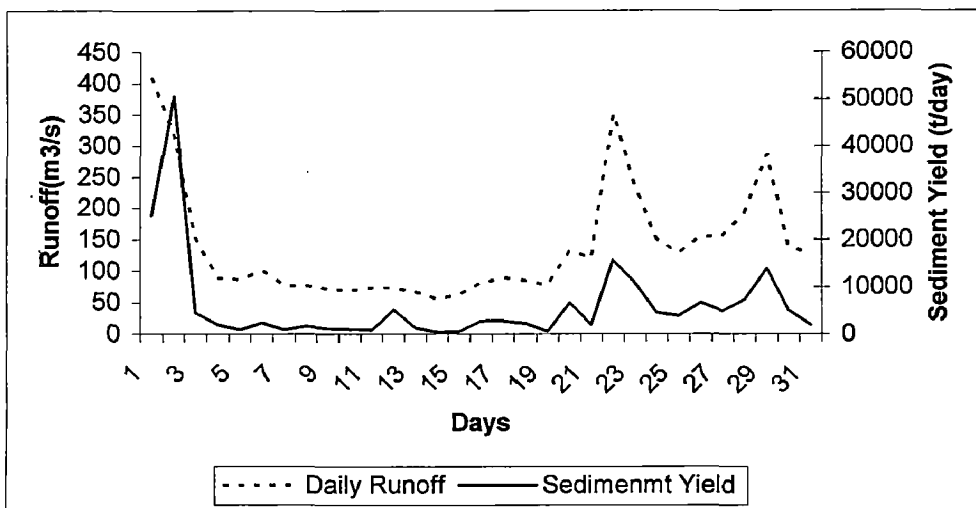
**Fig. C5 Daily runoff recorded at Mainachuli**

## APPENDIX – D      DAILY SEDIMENT YIELD AND RUNOFF DATA FROM 2001 TO 2003

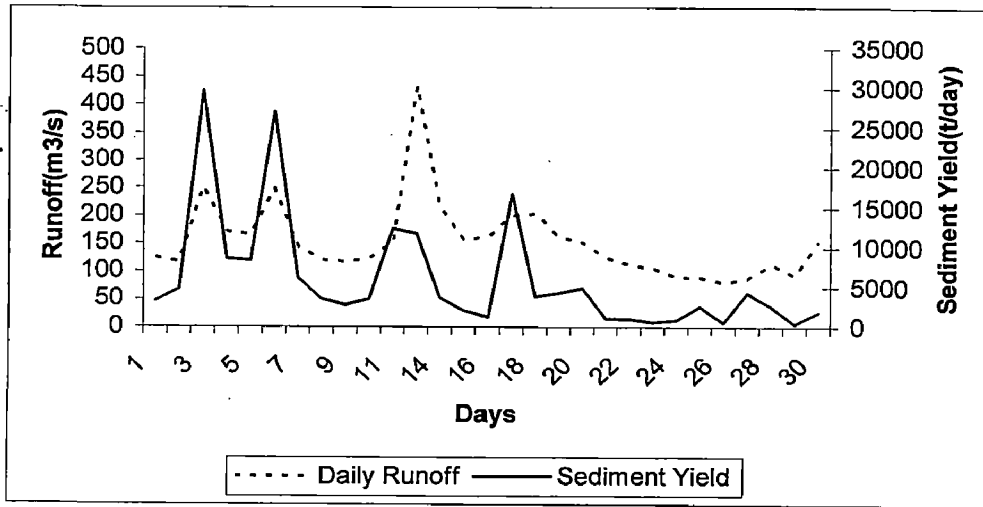
---



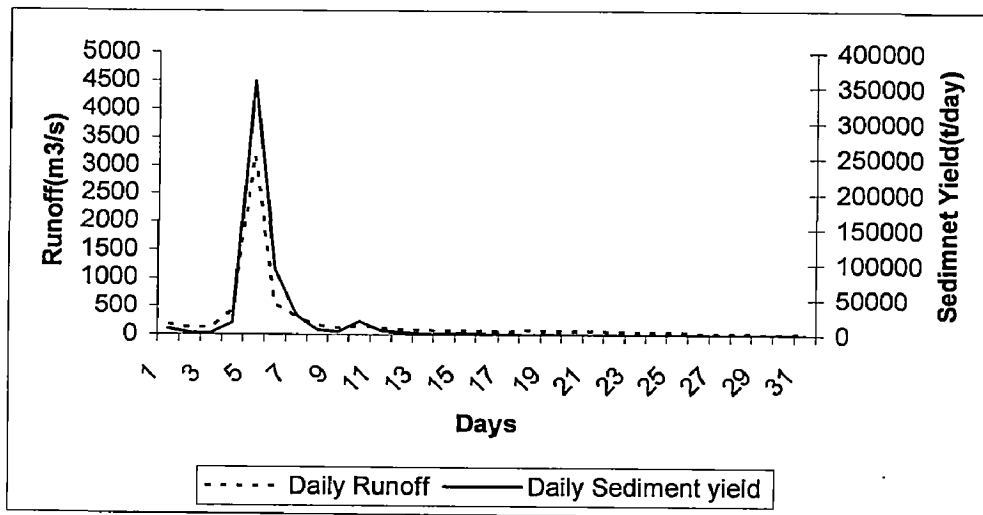
**Fig. D1 Daily sediment yield and Runoff data at Mainachuli ( July, 2001)**



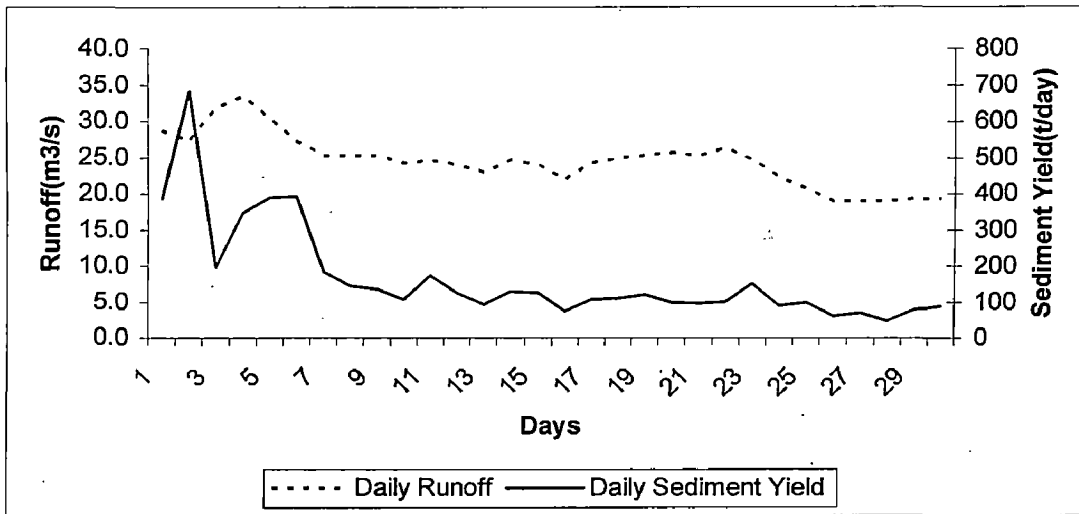
**Fig. D2 Daily sediment yield and runoff data at Mainachuli ( Aug., 2001)**



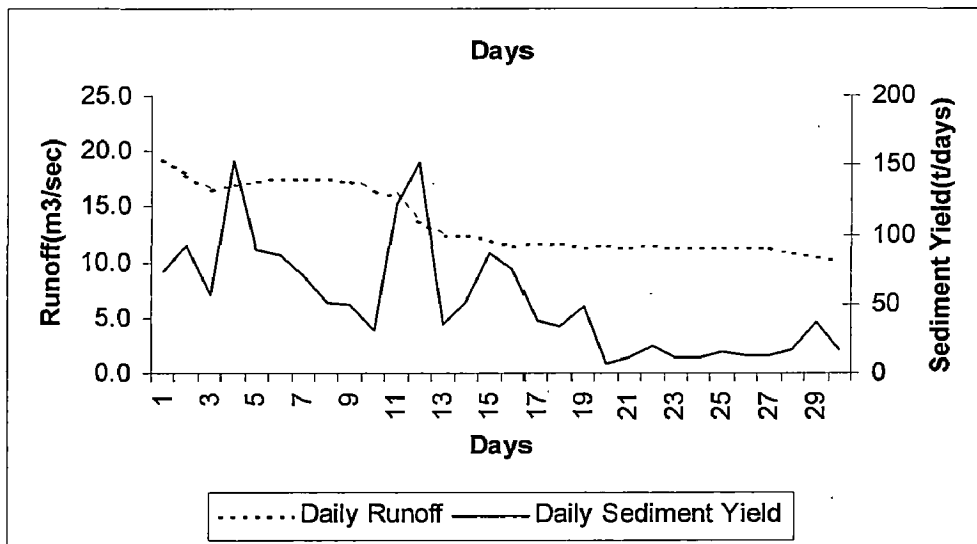
**Fig. D3 Daily sediment yield and runoff data at Mainachuli ( Sep., 2001)**



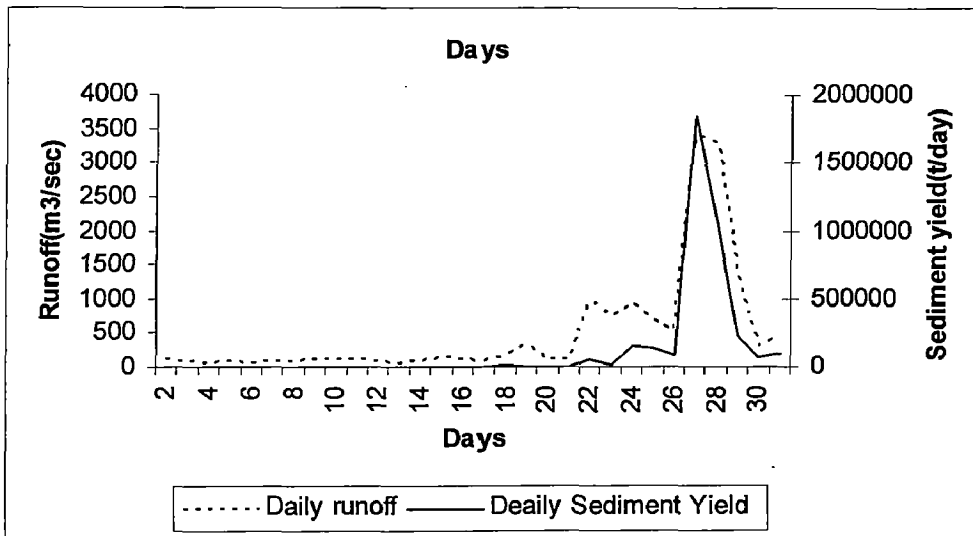
**Fig. D4 Daily sediment yield and runoff data at Mainachuli ( Oct., 2001)**



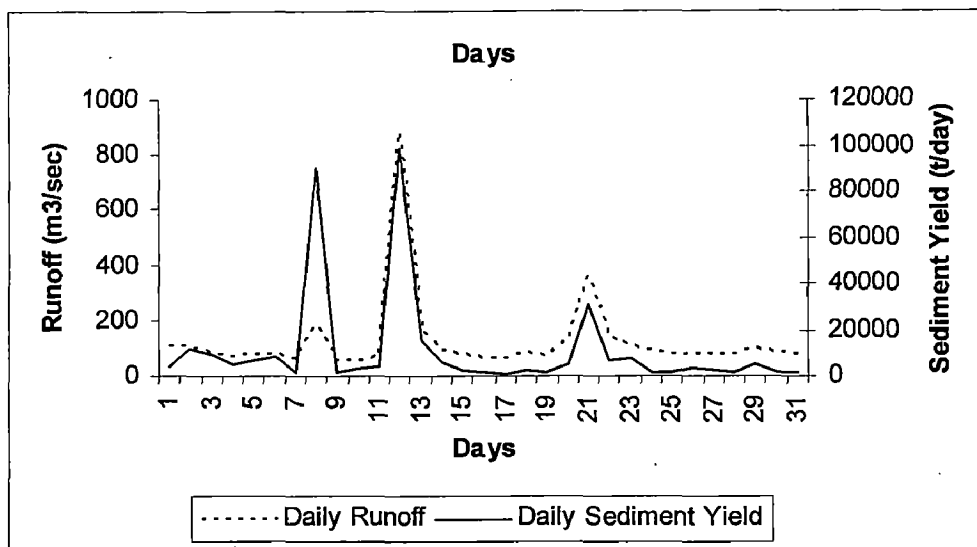
**Fig. D5 Daily sediment yield and runoff data at Mainachuli ( Nov., 2001)**



**Fig. D6 Daily sediment yield and runoff data at Mainachuli ( Dec., 2001)**



**Fig. D7 Daily sediment yield and runoff data at Mainachuli ( July, 2002)**



**Fig. D8 Daily sediment yield and runoff data at Mainachuli ( Aug., 2002)**

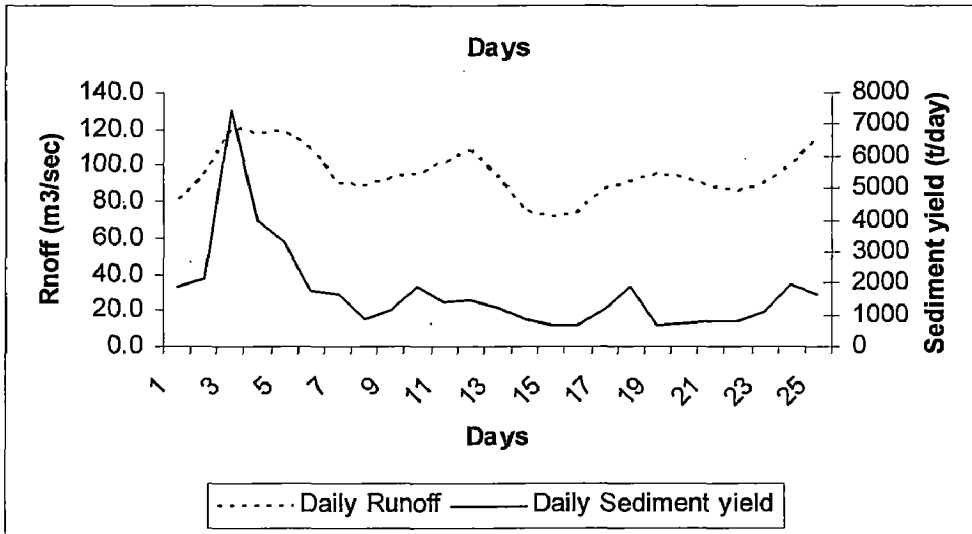


Fig. D9 Daily sediment yield and runoff data at Mainachuli ( Sep., 2001)

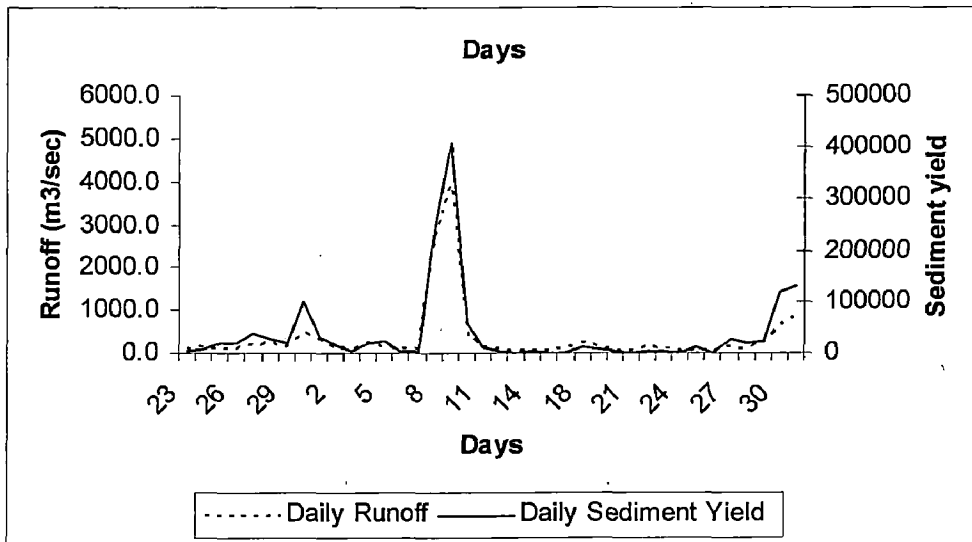


Fig. D10 Daily sediment yield and runoff data at Mainachuli ( July, 2003)



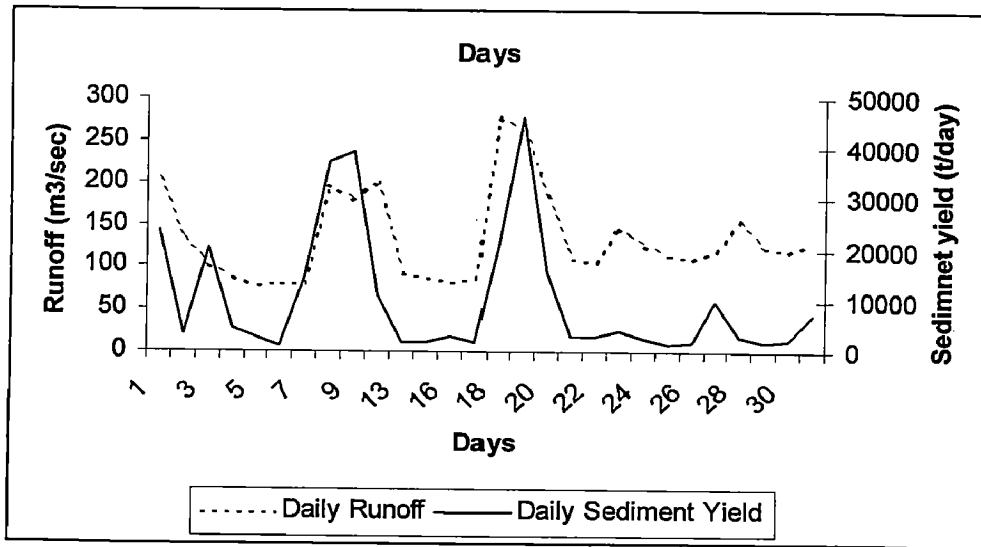


Fig. D11 Daily sediment yield and runoff data at Mainachuli ( Aug., 2003)

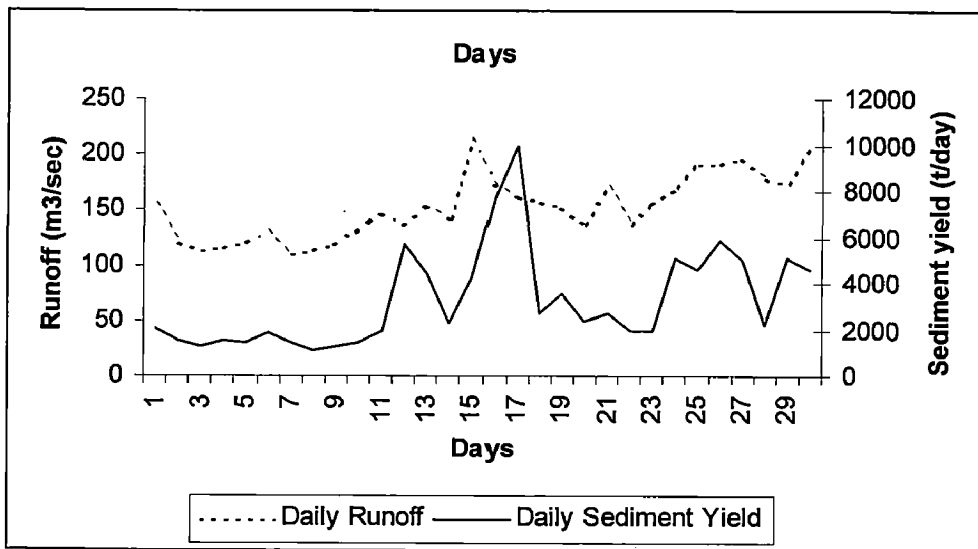
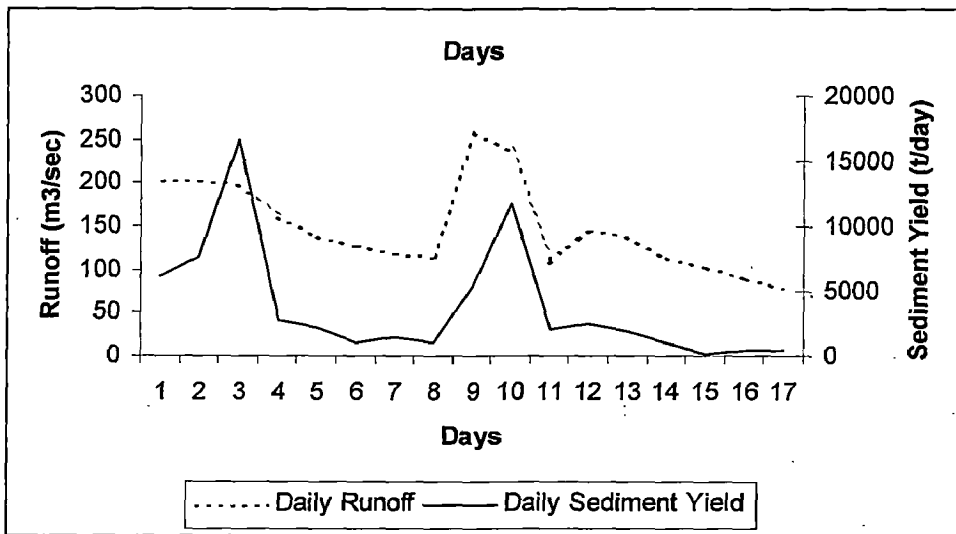


Fig. D12 Daily sediment yield and Runoff data at Mainachuli ( Sep., 2003)



**Fig. D13 Daily sediment yield and Runoff data at Mainachuli ( Oct., 2003)**